

PERMUTATION TESTS FOR RANDOMLY RIGHT CENSORED DATA
CONSISTING OF BOTH PAIRED AND UNPAIRED OBSERVATIONS

By

MICHAEL JAMES DALLAS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1997

© Copyright 1997

by

Michael James Dallas

To my wife Amy

ACKNOWLEDGEMENTS

I would like to thank Dr. P.V. Rao for his suggestions, guidance, optimism, and for all of the time he spent working with me during the course of my research. I am very fortunate to have had him as my major advisor. I would also like to thank Dr. Chang, Dr. Randles, Dr. Shuster, and Dr. W. Mendenhall (from the Department of Oncology) for serving as members of my supervisory committee, and Dr. S. Kramer (from the Department of Accounting) for attending my defense. In addition, I would like to express appreciation to Dr. Thomas O'Gorman from the University of Northern Illinois for sharing some Fortran code which proved to be useful in completing this work.

During my years as a graduate student in the Department of Statistics, I learned from each of my professors. Also, I was fortunate to be offered assistantship appointments through the Division of Biostatistics. While working for the Division of Biostatistics, I worked with several supervisors and I learned from each of them as I gained valuable practical consulting experience.

Finally, special thanks go to my wife Amy for her understanding and support throughout my graduate school experience.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT	vii
CHAPTERS	
1 INTRODUCTION	1
1.1 Censored Data	1
1.2 Overview	2
2 FORMULATION OF THE PROBLEM	5
2.1 Introduction	5
2.2 Notation and Assumptions	5
2.3 Literature Review	6
2.4 Objective	14
3 TWO TESTS OF H_0	16
3.1 Introduction	16
3.2 Permutation Tests	16
3.3 Rationale for Choosing Permutation Test Procedures	19
3.4 Rationale for the Choice of Test Statistics	21
3.5 Accommodating Right Censored Data	22
3.6 The Permutation Method	25
3.7 Overview of the D -PW and D -AK Procedures	30
4 SIMULATION STUDY	31
4.1 Introduction	31
4.2 Description	31
4.3 Results	39
4.4 Explanation of Results	60

5	APPLICATION AND DISCUSSION	61
5.1	Introduction	61
5.2	Analysis of Skin Graft Data	61
5.3	The Effect of Discarding Unpaired Observations	62
5.4	Extensions to the Research	70
5.5	Recommendations and Conclusions	71
APPENDICES		
A	ASYMPTOTIC DISTRIBUTIONS OF V_{PW} AND V_{AK}	72
B	FORTRAN CODE	78
C	SIMULATION RESULTS FOR THE $D-PW$ AND $D-AK$ TESTS ..	109
D	SIMULATION RESULTS FOR THE PAIRED PW AND AK TESTS	143
REFERENCES		145
BIOGRAPHICAL SKETCH		149

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

PERMUTATION TESTS FOR RANDOMLY RIGHT CENSORED DATA
CONSISTING OF BOTH PAIRED AND UNPAIRED OBSERVATIONS

By

Michael James Dallas

December 1997

Chairman: Pejaver V. Rao
Major department: Statistics

Two test procedures are proposed for comparing two survival distributions on the basis of randomly right censored data consisting of both paired and unpaired observations. Our procedures generalize the pooled rank test statistic of Hollander, Pledger, and Lin. One generalization adapts the Prentice-Wilcoxon scores, while the other adapts the Akritas scores. The use of these particular scoring systems in pooled rank tests with randomly right censored paired data has been advocated by O'Brien and Fleming, Dabrowska, Akritas, and Woolson and O'Gorman.

Although we provide asymptotic distributions of the test statistics by using counting process techniques, our test procedures utilize the permutation distributions of the test statistics based on a novel manner of permuting the scores. Permutation test methodology is a classical idea first suggested by Fisher. Such tests are attractive since they require only mild assumptions regarding the observations for valid use, are distribution-free (conditional on the observed data), and do not rely on asymptotics.

Because the tests are computationally intensive, we develop and provide computer programs which are key contributions, essential for practical implementation of the

tests. Permutation versions of tests for right censored paired data and for two independent right censored samples which use the proposed scoring systems are obtained as special cases of our test procedures.

Simulation results show that our test procedures are generally comparable to each other in terms of power. The nominal levels are basically maintained by the tests under identical exponential and log-logistic distributions for the survival times, which also include an appropriate term added to provide correlation. The tests have good power for detecting scale and location shifts in the exponential and log-logistic distributions. Weibull alternatives to the exponential distribution can be detected with moderate power provided the sample sizes are large and amount of censoring is low.

The tests are illustrated with a real data set. We also demonstrate the advantages of our test procedures in terms of including randomly occurring unpaired observations, which are discarded in test procedures for paired data.

CHAPTER 1 INTRODUCTION

1.1 Censored Data

Clinical studies are often conducted to compare medications or treatment methods with regard to length of time of effectiveness or length of time required to cure disease. Analysis of time data resulting from such studies is called survival analysis. Survival time is the time until the event of interest occurs.

Sometimes a survival time may be censored. If the event of interest has not occurred before the end of the observation time period, then the survival time is said to be right censored. The observation time period for a subject begins when the subject enters the study and ends when the subject exits the study or when the study terminates.

Right censoring can be broadly classified into three categories. Type I censoring occurs when the observation period for each subject is preset to be a certain length of time, common to all subjects. Survival times exceeding the observation period would be right censored. Type II censoring occurs when the study ends as soon as the r th event is observed, where r is a predetermined integer less than N , the total number of subjects in the study. Thus, the $N - r$ unobserved survival times would be right censored. The third type of censoring is called random censoring, which occurs when each subject's observation period is a random variable. Random censoring typically occurs in clinical trials, where subjects are entered and removed from the study at different times. Survival times exceeding observation periods would be right censored.

Survival times can also be left censored. Subjects who experience the event of interest prior to their initial observation of the study yield left censored survival times. Although less common than right censoring, left censoring sometimes occurs in clinical trials. In this dissertation, however, it is assumed that all censoring is right censoring.

It should be noted that survival analysis can be used when responses are not measured on a time scale. For example, in a study of hospital costs incurred by patients, total hospital cost to a patient may be regarded as a survival time. Uncensored responses would be obtained from patients whose total costs are known, whereas censored responses would be obtained from patients for whom we know only that their total costs exceed certain amounts. The patients who are still hospitalized and accruing cost at the time the costs are recorded would yield right censored responses.

1.2 Overview

This dissertation is concerned with the problem of testing the equality of two survival distributions using randomly right censored survival time data consisting of independent pairs of observations from a bivariate distribution, as well as independent observations from one or both of the marginal distributions.

As noted by Kalbfleisch and Prentice (1980), survival times may occur in pairs either naturally, for example, right and left eye observations, or by experimental design, when subjects are paired based on one or several characteristics. In addition to the paired survival times, a data set may also include independent unpaired survival times that provide information about only the marginal distributions. These unpaired observations may occur either randomly or by experimental design.

Data sets containing both paired and unpaired survival times will arise: (1) in survival studies employing matched-pair designs, when for some pairs, one of the two members of the pair drops out of the study before any observations are made on the

pair, while the other member of the pair continues on the study; (2) in survival studies employing matched-pair designs or for naturally paired survival time observations, when for some pairs, only an observation on one member of the pair is available due to circumstances beyond the control of the observer; (3) in survival studies designed in such a way that some subjects yield observations from a bivariate distribution, while other subjects yield single observations from a marginal distribution, as determined by physical conditions of the subjects.

Areas of study that may yield randomly right censored data containing both paired and unpaired observations include ophthalmology, renal disease, twin-studies, and nonmedical studies such as those involving automobiles or airplanes. The following are some examples where randomly right censored data consisting of both paired and unpaired observations will arise.

- Cancer studies to compare survival times under a standard therapy to those under an experimental therapy using a matched-pair design. When the data are recorded, it is realized that due to technical difficulties in the data collection process, not all of the observations are paired; some “pairs” yield only a standard therapy observation, and some “pairs” yield only an experimental therapy observation. The usual objective in such studies is to determine whether the experimental therapy has a longer survival time compared to the standard therapy.
- Studies to compare survival times of two competing types of hip [knee] replacements. Subjects enrolled in the study will either have one or two damaged hips [knees], and will be assigned either one (at random) or both types respectively. The usual objective in such studies is to determine which type has the longer effectiveness time.

As a specific example, Batchelor and Hackett (1970) reported the survival times of closely and poorly human lymphocyte antigen (HLA) matched skin grafts on the same burned individual. The data set consists of eleven cases where both a poorly and closely matched observation were recorded, four cases where only a poorly matched observation was recorded, and one case where only a closely matched observation was recorded. The five unpaired cases were a result of technical difficulties. The objective of this study was to determine whether survival time of closely matched grafts was longer than that of poorly matched grafts.

In this dissertation, following the convention employed by numerous researchers, it is assumed that members of the same pair have equal censoring times. When survival times are recorded on the same subject, this assumption is quite natural, but when pairs are matched through experimental design, censoring time can differ for the two members of the pair. In such a case, the minimum of the two censoring times is taken to be the common censoring time for the pair.

Chapter 2 provides notation and assumptions, a brief literature review, and the general objective of the dissertation. Chapter 3 contains the description and construction of the two test procedures introduced in this dissertation. Results of simulation studies investigating the performance of the new test procedures are given in Chapter 4. Chapter 5 presents a real data example to illustrate the use of these tests. Gains in power over tests that use just the paired observations, possible extensions to the research, recommendations for use of the tests, and conclusions of the dissertation are also given in Chapter 5.

CHAPTER 2 FORMULATION OF THE PROBLEM

2.1 Introduction

Section 2.2 presents notation and assumptions that will be used throughout the dissertation. A brief literature review is given in Section 2.3, while Section 2.4 provides a description of the specific objective of this dissertation.

2.2 Notation and Assumptions

Let $(X'_i, Y'_i, C_i), (X'_{n+j}, C_{n+j}), (Y'_{n+k}, C_{n+s+k}), i = 1, 2, \dots, n; j = 1, 2, \dots, s; k = 1, 2, \dots, t;$ be independent random vectors, where the X'_i and Y'_i represent survival times, and the C_i represent censoring times. Let (X'_i, Y'_i, C_i) be distributed as (X', Y', C) ; let (X'_{n+j}, C_{n+j}) be distributed as (X', C) ; and let (Y'_{n+k}, C_{n+s+k}) be distributed as (Y', C) , where X', Y' , and C are continuous random variables.

The observed survival times are $X = \min(X', C)$ and $Y = \min(Y', C)$. In addition to X and Y , we also observe $\delta_x = I(X' \leq C)$ and $\delta_y = I(Y' \leq C)$, where $I(\cdot)$ is the indicator function. Thus, δ_x and δ_y specify whether an observed survival time is a censored or a true (uncensored) survival time. The observed data can be expressed as $(X_i, \delta_{x_i}, Y_i, \delta_{y_i}), (X_{n+j}, \delta_{x_{n+j}}), (Y_{n+k}, \delta_{y_{n+k}}), i = 1, 2, \dots, n; j = 1, 2, \dots, s; k = 1, 2, \dots, t.$

The following assumptions will be used throughout this dissertation. The first two assumptions involve the variables X', Y' , and C , while the third assumption concerns the nature of randomly occurring unpaired observations.

Assumption 1: (X', Y') and C are independent.

Assumption 2: X' has survival function $S_{X'}(\cdot) = P(X' > \cdot)$, and Y' has survival function $S_{Y'}(\cdot) = P(Y' > \cdot)$.

Assumption 3: Randomly occurring unpaired observations are not caused by any reasons associated with survival.

Assumption 3 implies that the process that causes randomly occurring unpaired observations is ignorable for inferential purposes. Additional assumptions and notation will be introduced as needed.

There are three practically meaningful cases corresponding to different combinations of values of n , s , and t :

Case 1: $n = 0, s > 0, t > 0$

Case 2: $n > 0, s = 0, t = 0$

Case 3: $n > 0, \max(s, t) > 0$

The tests developed in this dissertation can be applied to any of the above cases. For right censored data arising from Case 1 or from Case 2, there exist several nonparametric tests which will be discussed in Section 2.3. Thus the two tests developed in this dissertation provide new alternatives for handling these cases. There do not exist any established nonparametric tests for right censored data arising from Case 3. Consequently, we focus on the use of the new tests for Case 3.

2.3 Literature Review

There exist a number of standard nonparametric tests for comparing two independent ($n = 0, s > 0, t > 0$) right censored samples. Such tests include the generalized Wilcoxon tests of Gehan (1965) and of Peto and Peto (1972), and the log rank test developed by Mantel (1966).

Nonparametric tests for analyzing paired ($n > 0$, $s = 0$, $t = 0$) right censored data have been studied fairly extensively. These tests can be broadly classified into two categories: (1) tests based on ranks of within pair differences and (2) tests based on pooled ranks. In discussing the details of the paired tests, the following concepts are needed: a pair is said to be doubly censored if both of the survival times of the pair are censored; singly censored if one of the survival times of the pair is censored; uncensored if none of the survival times of the pair is censored.

Among the numerous researchers who have studied within pair procedures are Woolson and Lachenbruch (1980), Popovich and Rao (1985), Dabrowska (1990), and Raychaudhuri and Rao (1996).

Woolson and Lachenbruch (1980) based their test on the pair differences, $D_i = X_i - Y_i$, under the assumption that the random variable $D' = X' - Y'$ is symmetrically distributed about a location parameter θ . The authors developed a score test of $H_0 : \theta = 0$ vs. $H_a : \theta > 0$, where the scores are based on the generalized rank vector (Prentice, 1978) of $|D|$. The derivation of the Woolson-Lachenbruch test follows that of the locally most powerful signed-rank test in the uncensored case.

The test statistic has the form

$$T_{WL} = T_u + T_c,$$

where T_u and T_c are signed-rank test statistics based on uncensored and singly censored pairs respectively. Woolson and Lachenbruch illustrated the use of T_{WL} assuming the loglinear model:

$$\begin{aligned} \log X'_i &= \theta + \eta_{1i} + \alpha_i, \\ \log Y'_i &= \eta_{2i} + \alpha_i \end{aligned} \tag{2.1}$$

where η_{1i} and η_{2i} are mutually independent random variables from a common distribution, and α_i is a "pairing" variable, distributed independently of η_{1i} and η_{2i} . A

logistic form for the density of $(\log X'_i - \log Y'_i)$ yields the censored data version of the signed Wilcoxon test, while a double exponential form yields the censored data version of the sign test. The authors suggested that for small sample sizes, a test based on the permutation distribution of T_{WL} can be used, while for large sample sizes a normal approximation will be appropriate. Asymptotic properties of tests based on T_{WL} were later thoroughly investigated by Dabrowska (1990).

Popovich and Rao (1985) developed an alternative to T_{WL} . The Popovich and Rao test is also based on the random variable D' , but their statistic has the form

$$T_{PR} = L_u T_u + L_c T_c. \quad (2.2)$$

Here T_u is a score statistic based on the signed-ranks of differences involving uncensored pairs, and T_c is the sign statistic based on singly censored pairs.

The authors showed that T_{PR} has several desirable properties including computational simplicity, asymptotic normality (under certain regularity conditions), and distribution-freeness (conditional on the number of uncensored and singly censored pairs, when T_u is conditionally distribution-free). Popovich and Rao showed that if L_u and L_c are functions of uncensored and singly censored pairs, then T_u and T_c are independent, conditional on the number of uncensored and singly censored pairs. This fact simplifies construction of the null distribution of T_{PR} .

Popovich and Rao used a simulation study to investigate the properties of tests based on T_{PR} in the special case where T_u is the Wilcoxon signed-rank statistic. In this special case, the exact critical values for T_{PR} can be easily determined from the widely available critical values of the Wilcoxon signed-rank and sign statistics. In the simulation study, four different distributions were considered for $(\log X'_i - \log Y'_i)$ in the loglinear model (2.1). Also, four different choices for the coefficients (weights) were considered: (1) $L_u^2 = L_c^2 = 0.5$, (2) L_c^2 proportional to the number of singly censored pairs, (3) L_c proportional to the number of singly censored pairs, and (4)

L_c proportional to the null standard deviation of T_u . The authors concluded that, in general, the power of the test based on T_{PR} is at least as high as the power of the test based on T_{WL} . They also concluded that none of the four weights emerged as the best in general, but they recommended weight system (2) if a single choice is to be made.

Dabrowska (1990) considered the problem of testing the null hypothesis that (X', Y') is equal in distribution to (Y', X') . She used a counting process representation to express a general form of the censored data version of the signed-rank test statistic,

$$T_D = \int_0^\infty K_u(v) d(N_1(v) - N_2(v)) + \int_0^\infty K_c(v) d(N_3(v) - N_4(v)),$$

where K_u and K_c are scoring systems and the N_k , $k = 1, 2, 3, 4$, are counting processes that count occurrences of uncensored and singly censored pairs within classes defined by the sign of within pair differences. By using a specific scoring system in T_D , the test statistic of Woolson and Lachenbruch (1980), T_{WL} , can be obtained as a special case.

Dabrowska showed that, under the null hypothesis and regularity conditions, $n^{-1/2}T_D$ converges in distribution to a mean 0 normal random variable. Specifically, both the signed Wilcoxon and sign forms of T_{WL} have this convergence property. The author demonstrated that efficacies of tests based on $n^{-1/2}T_D$ can be derived under contiguous alternatives by employing Le Cam's third lemma (Hájek and Šidák, 1967).

Dabrowska calculated the asymptotic relative efficiency (ARE) of the test based on T_D with respect to the likelihood ratio test based on the within pair absolute differences, the signs of the differences, and censoring indicators. She concluded that the censored data versions of the signed-rank tests are generally inefficient due to inappropriate scores assigned to the uncensored and singly censored observations and due to the fact that doubly censored observations are not informative. This fact was

illustrated with loglinear models parameterized as in (2.1). An ARE comparison was made among the censored data versions of the sign, signed Wilcoxon, and normal scores tests, which demonstrated that among these tests, either the signed Wilcoxon or the normal scores test should be used.

Raychaudhuri and Rao (1996) used counting process representations to study the efficacies of the tests based on T_{WL} and T_{PR} . The efficacies were obtained under a contiguous sequence of parametric alternatives. To compare the two tests, the authors expressed T_{WL} in the form of T_{PR} ; that is, both test statistics were expressed in the form

$$T_{RR} = L_u T_u + L_c T_c,$$

as in (2.2). Based on efficacies, the optimal coefficients for the two tests were obtained.

Asymptotic relative efficiency comparisons were made under the loglinear model (2.1) assuming a variety of values for the probability of double censoring and for the correlation coefficient between the log of the paired survival times. Three forms of T_{RR} , two forms of T_{WL} and the usual form of T_{PR} , were compared with regard to ARE. The two forms of T_{WL} considered were formed by taking (1) the censored data version of the sign statistic, and (2) the censored data version of the Wilcoxon signed-rank statistic. The usual form of T_{PR} was used by taking T_u to be the Wilcoxon signed-rank statistic based on uncensored pairs and T_c to be the sign statistic based on singly censored pairs. For comparison purposes three linear combinations—one with equal weights (weight system (1) from Popovich and Rao (1985)), one with weights proportional to the null standard deviations (weight system (4) from Popovich and Rao (1985)), and one with optimal weights—were formed for each of the three statistics above. The authors concluded that the tests based on T_{PR} perform as well as the tests based on T_{WL} , and that the tests based on T_{PR} perform better under heavy double censoring.

Among the numerous authors who have researched pooled rank procedures are Cheng (1984), O'Brien and Fleming (1987), Dabrowska (1990), and Akritas (1992).

Cheng (1984) introduced a class of pooled rank test statistics to test the null hypothesis that X' and Y' have equal distributions. The author utilized the test statistic

$$T_{CH} = n^{-2} \sum_{i=1}^n \sum_{i'=1}^n \psi(X_i, \delta_{x_i}, Y_{i'}, \delta_{y_{i'}}),$$

where ψ is a specific type of scoring function based on pooled ranks of the (X_i, δ_{x_i}) and (Y_i, δ_{y_i}) . Cheng showed that, under the null hypothesis and suitable regularity conditions, T_{CH} converges in distribution to a mean 0 normal random variable with a variance that can be consistently estimated. Related to this class of statistics are the class of statistics proposed by O'Brien and Fleming (1987) and the statistic proposed by Akritas (1992).

O'Brien and Fleming (1987) developed a test of the null hypothesis that X' and Y' have equal distributions using the rank transformation idea of Conover and Iman (1981) and Prentice-Wilcoxon scores (Prentice, 1978). Their class of test statistics is based on the differences $\Delta_i = k_{x_i} - k_{y_i}$, where (k_{x_i}, k_{y_i}) are appropriately chosen scores based on the pooled ranks of (X_i, δ_{x_i}) and (Y_i, δ_{y_i}) . Conditional on the scores assigned to the i th pair, and under the null hypothesis, k_{x_i} is equally likely to be the score assigned to (X_i, δ_{x_i}) and (Y_i, δ_{y_i}) ; therefore, the Δ_i are independently distributed with mean 0 and variance Δ_i^2 . The authors conjectured that under the Lindeberg Condition (Feller, 1966), the statistic

$$T_{OF} = \sum_{i=1}^n \Delta_i / (\sum_{i=1}^n \Delta_i^2)^{1/2}$$

converges in distribution to a standard normal random variable.

On the basis of a comparison of several scoring systems for defining the Δ_i , O'Brien and Fleming advocated using Prentice-Wilcoxon scores. To compute the resulting

statistic, denoted by T_{PW} , all uncensored times are ranked from smallest to largest, and n_l is defined as the number at risk in the pooled sample at the l th smallest uncensored time. With $s_m = \prod_{l=1}^m n_l / (n_l + 1)$, the m th smallest uncensored observation is assigned a score of $(1 - 2s_m)$, while each censored observation in the interval $[m, m+1)$ is assigned a score of $(1 - s_m)$. The authors recommended assigning tied uncensored times their average score. O'Brien and Fleming also suggested that for small sample sizes, a permutation test would be more appropriate than the normal approximation.

Dabrowska (1990) examined two pooled rank tests: the test based on T_{PW} and the pooled rank test using log rank scores, both of which were considered by O'Brien and Fleming (1987). Similar to the within pair tests she considered, efficacies were obtained. An ARE comparison relative to the fully efficient pooled rank test showed that the test based on T_{PW} had higher ARE than the paired log rank test under heavy censoring and/or high correlation between the survival times.

Akritis (1992), like O'Brien and Fleming (1987), developed a test based on the rank transformation procedure. Essentially, in this test, the observations are replaced by scores based on their pooled rank, and then a paired t-test is performed on these scores.

The Akritis scoring procedure will first compute the Kaplan-Meier (1958) survival function estimates, $\hat{S}_X(\cdot)$ and $\hat{S}_Y(\cdot)$, from the X and Y samples separately. Then the average estimate,

$$\bar{\hat{S}}(\cdot) = \frac{(\hat{S}_X(\cdot) + \hat{S}_Y(\cdot))}{2} \quad (2.3)$$

is computed, and each uncensored observation is assigned a score of $2n(1 - \bar{\hat{S}}(\cdot))$, while each censored observation is assigned a score of $2n(1 - \frac{1}{2}\bar{\hat{S}}(\cdot))$. These scores are then used in a paired t-test. The Akritis statistic will be denoted by T_{AK} .

Akritas proved a theorem stating that replacing the observations with the above scores does not change the asymptotic distribution of the paired t -test statistic. He performed a small simulation study to compare the finite sample size performances of tests based on T_{PW} and T_{AK} , assuming gamma distributions for the survival times. His comparison showed that the two procedures maintained their levels and had comparable power under both location and scale shift alternatives.

Within pair procedures, for uncensored data, are the conventional (signed-rank) procedures. The theory of pooled rank procedures will be discussed in Section 3.3. An important difference between the two types of procedures is that pooled rank tests utilize interpair information, which is ignored in within pair procedures. In fact, O'Brien and Fleming (1987) noted that the test based on their pooled rank statistic, T_{PW} , appropriately uses interpair information based on the degree of correlation; in uncorrelated data full use is made of interpair information, whereas as the correlation approaches 1, the use of interpair information decreases to zero. Albers (1988) noted that the use of interpair information can lead to efficiency advantages of pooled rank tests over within pair tests.

Comparisons between the two types of procedures with censored paired data were made by several authors. O'Brien and Fleming's (1987) comparison between the tests based on T_{PW} and T_{WL} via Monte Carlo simulation revealed that the test based on T_{PW} has power similar to or greater than the test based on T_{WL} under a variety of alternatives to equal exponential distributions. Based on ARE comparisons, Dabrowska (1990) recommended that under heavy censoring, the pooled rank tests she considered should be used in lieu of the censored data version of the signed Wilcoxon test. Woolson and O'Gorman (1992) used Monte Carlo methods to evaluate powers of several tests for censored paired data, including tests based on T_{WL} , T_{PW} , and T_{AK} , under eight different sets of assumptions regarding the distributions of the survival times. Their study revealed that (1) each test maintains its level, (2) tests based on

T_{AK} and T_{PW} have desirable power properties compared to the other tests, and (3) tests based on T_{AK} and T_{PW} have nearly equal powers in all eight situations. The authors also pointed out that T_{PW} and T_{AK} are similar algebraically and that the real difference between them is in the scoring systems. As a general conclusion, they recommended using either a test based on T_{PW} or T_{AK} .

Based on the aforementioned literature review, we can generally conclude that for censored paired data, a pooled rank test employing either the PW or AK scoring system should be used.

2.4 Objective

Several nonparametric tests for comparing the distributions of X' and Y' for unpaired (two independent sample) right censored data and also for paired right censored data are discussed above. In this dissertation we want to investigate nonparametric tests for right censored data consisting of both paired and unpaired observations. The key consideration in developing such tests is to utilize both paired and unpaired observations. An extensive literature review has revealed that this problem has yet to be thoroughly investigated. The problem is briefly addressed by Cheng (1984) as an extension to the paired data problem.

The general objective of this dissertation is to develop tests of

$$H_0 : S_{X'}(\cdot) = S_{Y'}(\cdot), \quad (2.4)$$

using both paired and unpaired observations, and to study the properties of these tests. Specifically we focus on permutation tests, based on reasons to be discussed later. To achieve our objective, the following specific tasks have been undertaken:

- 1) Develop appropriate test statistics.

- 2) Formulate test procedures based on the permutation distributions of the test statistics.
- 3) Develop computing routines necessary to execute the tests.
- 4) Evaluate power of the proposed tests through simulation study.
- 5) Demonstrate the usefulness of the tests.

CHAPTER 3 TWO TESTS OF H_0

3.1 Introduction

In this chapter, two permutation tests for testing H_0 at (2.4)— X' and Y' have identical distributions—are formulated. Permutation tests are introduced in Section 3.2. Section 3.3 discusses the rationale for choosing the permutation test approach for our test of H_0 . Section 3.4 discusses the rationale for the proposed test statistics, which are then described in detail in Section 3.5. In Section 3.6, the specific method of implementing the two tests is described. A brief summary of the two test procedures is given in Section 3.7.

3.2 Permutation Tests

Edgington (1987) defines a permutation test as a statistical test for which the significance of experimental results is determined by permuting the data repeatedly to compute the test statistic. In experimental settings, when treatments are randomly assigned to experimental units, the random assignment often forms the basis for permuting data. When this occurs, permutation tests are also referred to as randomization tests.

The steps necessary to execute a permutation test are as follows:

- 1) Determine an appropriate test statistic, T .
- 2) Compute the value of T , say T_{obs} , for the observed data.

- 3) Permute the data repeatedly, in such a way that, conditional on the observed data and under the null hypothesis, all permutations are equally likely. The set of all such permutations is called the reference set.
- 4) Compute the value of T , say T^* , for each permutation in the reference set.
- 5) Calculate the p-value as the proportion of T^* values at least as favorable to rejection of H_0 as T_{obs} .

As pointed out by Edgington (1987), permutation tests are not alternatives to conventional tests; they are tests in which only the significance is determined by the permutation procedure. Moreover, Noreen (1989) remarked that if significance can be assessed using conventional techniques, then significance can almost always be assessed using permutational techniques, but the converse is not true, and the p-values from the two approaches can be considerably different for some data sets. Bradley (1968) pointed out that certain "eminent" statisticians including Sir Ronald Fisher and Oscar Kempthorne believe that the permutation test is the truly correct test to assess significance.

Permutation tests have several advantages. Mainly, the tests require minimal assumptions concerning the observed data. In experimental settings, random assignment to the experimental units is the only assumption necessary for valid use of the test. Assumptions such as random sampling (which rarely, if ever, occurs in experimentation) and underlying normality (or other parametric distribution) are unnecessary. Secondly, any statistical test is a distribution-free test when significance is determined by the permutation procedure. Also, permutation tests do not rely on asymptotic theory for determining significance.

Permutation tests require a great amount of computing, which was an obvious drawback to their use until recently. Presently, however, executing permutation tests is quite feasible due to advances in computer technology along with advances in fast

and efficient computing algorithms. Even with high-speed computing, tests based on all possible permutations of the data may not be feasible in certain situations, but Dwass (1957), followed by others, helped to resolve this problem by studying tests based on a random sample of all possible permutations. These researchers observed that although they are less powerful than tests based on all possible permutations, tests based on a random sample of all possible permutations are valid, and as the size of the random sample increases, the power of the test approaches the power of the test based on all possible permutations. Dwass (1957) and Recchia and Rocchetti (1982) have shown that a random sample size of 1000 provides power comparable to the power of the test based on all possible permutations, regardless of the number of all possible permutations. Thus, as stated by Pagano and Tritchler (1983), the power loss incurred by using a random sample usually is not great enough to warrant the computational burden of complete enumeration. Tests based on all possible permutations are called exact tests, while tests based on a random sample are called approximate tests.

Fisher (1935) was the first to propose the idea of permutation tests. Theoretical properties of permutation tests were investigated by several researchers including Pitman (1937a, 1937b, 1938), Pearson (1937), Scheffe (1943), Noether (1949), Lehmann and Stein (1949), Hoeffding (1951, 1952), Puri and Sen (1971), and Lehmann (1986). Some of the theory is devoted to permutational limit theorems, which pertain to the convergence of the permutation distribution of the test statistic as the sample size increases. Such theorems were useful before the days of high-speed computing, when enumerating the permutation distribution was extremely laborious. Thus, limit theory may not be as useful or necessary in the modern era, where a great amount of research focuses on the actual use of the tests and the associated computing aspects. Edgington (1987) has provided the most comprehensive overview of the practical use of randomization tests for a variety of settings. Researchers who have studied computational aspects include Green (1977), Mehta, Patel, and Tsaitis (1984), Mehta,

Patel, and Wei (1988), Noreen (1989), and Oden (1991). Oden suggested that if computers were available in the early 1900s, classical normal-theory based tests might not have even been developed; instead research efforts may have been devoted to randomization and other computer-based methods.

3.3 Rationale for Choosing Permutation Test Procedures

Although Woolson and Lachenbruch (1980) and O'Brien and Fleming (1987) acknowledged the importance of exact test procedures, the test procedures mentioned in Section 2.3 are based on the asymptotic normal distributions of the test statistics. It is reasonable to question the validity and accuracy of asymptotic procedures for finite samples.

Asymptotic approximations are particularly questionable when there exist a number of tied survival times, which often result from the recording of survival times on the basis of periodic examinations of study subjects. Also, heavy censoring is a cause for concern with utilizing asymptotics (Prentice, 1978).

Despite questions regarding the validity of tests based on asymptotics with censored data, practical computing limitations of the past forced utilization of asymptotic procedures. Several examples of this situation follow. Popovich and Rao (1985) could not include crucial small sample comparisons between their proposed test procedure and that of Woolson and Lachenbruch (1980), because they noted that evaluating the permutation distribution was impractical due to "lack of computing facilities." Thus, one is forced to select a test procedure based on its asymptotic properties. Dabrowska (1989) considered the asymptotic distribution of her test statistic due to her belief that conditional (permutation) tests were not feasible. Fleming and Harrington (1991) claimed that it is impractical to compute exact finite sample distributions for censored data linear rank statistics; therefore, they utilized a counting process framework to

study asymptotic test procedures. Andersen and Borgan (1985), however, noted that the counting process approach cannot solve the finite sample size problem.

The claims of practical computing limitations are not valid in the present era. In fact, many researchers who realized the impact of high-speed computing developed exact (permutation) tests for censored data. These researchers include Padgett and Wei (1983), Edgington and Gore (1986), Good (1989), Janssen and Brenner (1991), and Sun and Sherman (1996). Padgett and Wei (1983) noted that test procedures of this type are highly desirable since they are distribution-free for any sample size. Following this recent trend, we focus on the use of permutation tests for testing H_0 .

In addition to the general advantages of permutation tests given in Section 3.2, advantages of permutation tests that are particularly useful for the problem we consider include the following. The test statistic needs only to include a component which reflects the difference in survival distributions. It is unnecessary to include an estimate of the variance of the given component, which would be required to normalize the test statistic. Since the tests can accommodate a variety of random assignment procedures, experimental design possibilities are expanded (Edgington, 1987). The design can be chosen based only on efficient use of experimental units, without concern for the complexity of the analysis. Thus, an experimental design that yields both paired and unpaired observations is a viable possibility. Since permutation tests are versatile procedures, they can be used to develop new tests which are useful in situations where no standard methodology exists (Edgington, 1987), as is the case for the problem addressed in this dissertation.

Although we assume throughout that X' , Y' , and C have continuous distributions, by using a permutation approach our procedures utilize an appropriate conditional null distribution regardless of whether there are ties among the survival times (or scores generated from the survival times). This is an advantage of permutation tests which is nicely illustrated in Randles and Wolfe (1991). Thus, without modification,

our procedures are applicable to situations where X' , Y' , and C have discrete distributions, as will be the case with grouped survival times which are common in survival studies.

3.4 Rationale for the Choice of Test Statistics

Recall that we observe $(X_i, \delta_{x_i}, Y_i, \delta_{y_i})$, $(X_{n+j}, \delta_{x_{n+j}})$, $(Y_{n+k}, \delta_{y_{n+k}})$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, s$; $k = 1, 2, \dots, t$. Using these data we need to select a statistic, V , to test $H_0 : S_{X'}(\cdot) = S_{Y'}(\cdot)$. Our choice of V is motivated by the following.

For paired ($n > 0, s = 0, t = 0$) uncensored data, Snijders (1981) studied pooled rank test procedures. Conditional on the observed configuration of ranks, Snijders developed the (conditionally) locally most powerful pooled rank test. He noted, however, that in general, evaluating the associated scores is difficult. Therefore, alternate scores have been advocated. Specifically, scores that are optimal for particular two-sample tests have been proposed for use in pooled rank tests with paired uncensored data (Snijders, 1981; Lam and Longnecker, 1983). This idea of ignoring the pairing while scoring the observations was also advocated by Conover and Iman (1981). The associated test statistics look like the standard two-sample rank test statistic, but the critical values are obtained by conditioning on the observed configuration of ranks. Essentially one uses the standard two-sample rank test statistic, but adjusts its variance to account for dependence.

Therefore, it is quite natural to extend this methodology to the case where unpaired observations are also realized ($n > 0, \max(s, t) > 0$). In fact, in such a case, Hollander, Pledger, and Lin (1974) advocated using the Mann-Whitney form of the two-sample Wilcoxon statistic,

$$U_1 = \sum_{i=1}^{n+s} \sum_{j=1}^{n+t} I(X_i > Y_j),$$

with the variance of U_1 adjusted to account for the pairing of the X_i and Y_i . U_1 belongs to a more general class of linear rank statistics given by

$$U_2 = \sum_{i=1}^{n+s} a(R_{x_i}),$$

where a is an appropriate score function for the two-sample case, and R_{x_i} is the rank of X_i in the pooled sample of $2n + s + t$ observations. In particular, when using the optimal score function for the two-sample location problem for the logistic distribution, $a(i) = i$, U_1 and U_2 are equivalent test statistics (see Randles and Wolfe, 1991).

We take U_2 as the starting point for defining the test statistic V . By utilizing the pooled ranks of the $2n+s+t$ observations, the test procedure based on V appropriately uses interblock information, a property of pooled rank tests which was discussed in Section 2.3. Additionally, we can construct a distribution-free exact test procedure based on V by conditioning on the observed configuration of ranks.

3.5 Accommodating Right Censored Data

In the presence of censoring, as with uncensored data, evaluating the scores for pooled rank tests with paired data is difficult. Therefore alternate scores have been advocated. Similar to the uncensored case, the scores or approximate scores (Prentice, 1978; Cuzick, 1985; Dabrowska, 1989, 1990; et al.) for particular two-sample tests with censored data have been proposed for use in pooled rank tests with paired censored data (O'Brien and Fleming, 1987; Albers, 1988; Dabrowska, 1989). Thus, following the uncensored data extension to the case where unpaired observations are also realized, we can generalize U_2 by considering test statistics of the form

$$V = \sum_{i=1}^{n+s} \dot{a}(\dot{R}_{x_i}, \delta_{x_i}).$$

Here \dot{a} is a two-sample score (or approximate score) function, and \dot{R}_{x_i} is the censored data rank, defined as in Prentice (1978), of X_i in the pooled sample of $2n + s + t$ observations. More precisely,

$$\dot{R}_{x_i} = \sum_{j=1}^{n+s} \delta_{x_j} I(X_j \leq X_i) + \sum_{k=1}^{n+t} \delta_{y_k} I(Y_k \leq X_i).$$

Thus, uncensored observations are ranked among themselves and each censored observation is assigned the same rank as the nearest uncensored observation less than or equal to it.

We will define two different scoring systems for V . We will use notation that helps to indicate the specific scores employed.

A natural scoring system follows from Prentice's (1978) censored data generalization of the Wilcoxon test. In this method, we pool the $2n + s + t$ observations into a single sample and assign the score $g_{\tilde{S}}(Z)$ to observation Z , where

$$g_{\tilde{S}}(Z) = \begin{cases} 1 - 2\tilde{S}(Z) & \text{if } Z \text{ is uncensored,} \\ 1 - \tilde{S}(Z) & \text{if } Z \text{ is censored.} \end{cases}$$

Here $\tilde{S}(Z) = \prod_{l=1}^Z \frac{n_l - d_l + 1}{n_l + 1}$, where n_l is the number at risk in the pooled sample of $2n + s + t$ observations at the l th smallest uncensored time, and d_l is the number of deaths at the l th smallest uncensored time. In the context of the two-sample problem, Gill (1980) showed that the above scores are optimal for detecting scale shifts in the log-logistic distribution of survival times, when $d_l = 1$ for all l (no tied uncensored times). In addition, our scoring system described here is the system used in the PW pooled rank procedure used with paired data, as described in Section 2.3, except for the way that tied uncensored times are handled. In our scoring system, ties are handled as in the Kaplan-Meier (1958) estimator. The resulting test statistic, V_{PW} , is the sum of the scores over the X sample:

$$V_{PW} = \sum_{i=1}^{n+s} g_{\bar{S}}(X_i).$$

The second scoring system follows from Akritas (1992), who utilized scores based on the redistribute mass to the right algorithm of Efron (1967). In this method, we pool the $2n + s + t$ observations into a single sample and assign the score $h_{\bar{S}}(Z)$ to observation Z , where

$$h_{\bar{S}}(Z) = \begin{cases} 1 - \bar{S}(Z) & \text{if } Z \text{ is uncensored,} \\ 1 - \frac{1}{2}\bar{S}(Z) & \text{if } Z \text{ is censored,} \end{cases}$$

and \bar{S} is the average Kaplan-Meier estimate given by (2.3). Akritas (1992) argued that by using the average survival estimate the above scores are related to the optimal log-logistic scores, which use the pooled sample survival estimate, for the two-sample scale problem. In addition, our scoring system described here is the system used in the *AK* pooled rank procedure used with paired data, as described in Section 2.3, except that the *AK* scoring system multiplies the scores by the total sample size. Clearly the sample size factor does not affect the p-value of the test, so this factor will not be utilized in our test procedure. The resulting test statistic, V_{AK} , is the sum of the scores over the X sample:

$$V_{AK} = \sum_{i=1}^{n+s} h_{\bar{S}}(X_i).$$

Two different test statistics for testing H_0 , V_{PW} and V_{AK} , in accordance with the respective scoring systems are now proposed. There do, of course, exist other scoring systems that could be used in V . However, we advocate the two systems described above based on (1) the desirable performance of the *PW* and *AK* scoring systems and the recommendations for their use in pooled rank tests for paired data, which

were noted in Section 2.3, and (2) the optimality property of the *PW* scoring system for unpaired data. Thus, it is reasonable to expect desirable properties for the scores we advocate with the combination of paired and unpaired observations.

By using counting process techniques, as in Fleming and Harrington, the null asymptotic distributions of the test statistics were obtained and provided in Appendix A. However, in order to complete our testing procedures, the permutation V^* values of these test statistics must be enumerated. The enumeration process is discussed in the following section.

3.6 The Permutation Method

Each of the two tests is based on the permutation principle, which is briefly described in Randles and Wolfe (1991), Section 11.1. For the problem under consideration, we are conditioning on the configuration of scores generated from the observed data, following Snijders (1981). Then, under H_0 and conditional on the scores for the n paired observations, there are 2^n equally likely arrangements of scores based on which of the 2 scores within each pair is regarded as the X sample score. Also, under H_0 and conditional on the scores for the $s + t$ unpaired observations, there are $\binom{s+t}{s}$ equally likely arrangements of scores based on which s of the scores are regarded as X sample scores. Thus, under H_0 there are

$$M = 2^n \times \binom{s+t}{s}$$

equally likely arrangements of scores.

Specifically, let $V_{PW_1}^*, V_{PW_2}^*, \dots, V_{PW_M}^*$ be the possible values for V_{PW} corresponding to the M arrangements of $g_{\hat{S}}(\cdot)$ scores. Then

$$P_{H_0}(V_{PW} = v | g_{\bar{S}}(o)) = \frac{\sum_{i=1}^M I(V_{PW_i}^* = v)}{M}, \quad (3.1)$$

where $P_{H_0}(\cdot | g_{\bar{S}}(o))$ stands for conditional (on observed $g_{\bar{S}}(\cdot)$ scores) probability calculated under the null hypothesis. Similarly,

$$P_{H_0}(V_{AK} = v | h_{\bar{S}}(o)) = \frac{\sum_{i=1}^M I(V_{AK_i}^* = v)}{M}. \quad (3.2)$$

The construction of the tests are based on (3.1) and (3.2) in the following manner. For an α -level test of H_0 , we reject H_0 if and only if the observed value of V_{PW} , $V_{PW_{obs}}$, $[V_{AK}, V_{AK_{obs}}]$ falls in a set R_α containing elements chosen from $\{V_{PW_1}^*, V_{PW_2}^*, \dots, V_{PW_M}^*\}$ $\{[V_{AK_1}^*, V_{AK_2}^*, \dots, V_{AK_M}^*]\}$ in a manner appropriate for the alternative hypothesis. Since the null distributions of V_{PW} and V_{AK} are discrete, there are a restricted number of possible α -levels. For example, a test that rejects H_0 for large values of V_{PW} $[V_{AK}]$ has M available α -levels if $\{V_{PW_1}^*, V_{PW_2}^*, \dots, V_{PW_M}^*\}$ $\{[V_{AK_1}^*, V_{AK_2}^*, \dots, V_{AK_M}^*]\}$ are distinct, and fewer than M available α -levels otherwise.

By employing the V_{PW} and V_{AK} test statistics and their associated null permutation distributions, we have introduced two procedures for testing H_0 . In what follows, the proposed procedures will be called the D - PW and D - AK procedures respectively.

We illustrate the details involved in conducting the D - PW and D - AK tests by considering a specific example.

Example 3.1 Suppose that we wish to test

$$H_0 : S_{X'}(\cdot) = S_{Y'}(\cdot) \text{ vs. } H_a : S_{X'}(\cdot) > S_{Y'}(\cdot),$$

using the data

Case Number						
	1	2	3	4	5	6
X	41	32	15 ⁺	35 ⁺	12	—
Y	22	40 ⁺	14	—	—	16

where a “+” indicates a censored observation.

Here we have $n = 3$, $s = 2$, and $t = 1$; so $M = 2^3 \times \binom{2+1}{2} = 24$. Since the alternative hypothesis implies that the X sample tends to have longer survival times than the Y sample, the critical region will contain an appropriate number of largest $V_{PW}^* [V_{AK}^*]$ values. For instance, (assuming 24 distinct $V_{PW}^* [V_{AK}^*]$ values) an $\alpha = 0.042 (= \frac{1}{24})$ level critical region, $R_{0.042}$ will contain the largest $V_{PW}^* [V_{AK}^*]$ value, whereas an $\alpha = 0.084 (= \frac{2}{24})$ level critical region, $R_{0.084}$ will contain the both second largest and largest $V_{PW}^* [V_{AK}^*]$ values. Also, the smallest level at which H_0 would be rejected, or the p-value, is the proportion of $V_{PW}^* [V_{AK}^*]$ values that are greater than or equal to the observed test statistic $V_{PW_{obs}} [V_{AK_{obs}}]$, in this case.

Tables 3.1 and 3.3 demonstrate the computation of scores for the V_{PW} and V_{AK} statistics respectively. Tables 3.2 and 3.4 detail the permutation methods used in calculating V_{PW}^* and V_{AK}^* values respectively.

Tables 3.1 and 3.2 illustrate the D - PW procedure.

Table 3.1. Computations for the V_{PW} statistic

X_i	δ_{x_i}	Y_i	δ_{y_i}	$\tilde{S}(X_i)$	$g_S(X_i)$	$\tilde{S}(Y_i)$	$g_S(Y_i)$
41	1	22	1	0.229	0.543	0.571	-0.143
32	1	40	0	0.457	0.086	0.457	0.543
15	0	14	1	0.800	0.200	0.800	-0.600
35	0	-	-	0.457	0.543	—	—
12	1	-	-	0.900	-0.800	—	—
-	-	16	1	—	—	0.686	-0.371

From Table 3.1, $V_{PW_{obs}} = \sum_{i=1}^5 g_S(X_i) = 0.572$. The 24 equally likely V_{PW}^* values are given in Table 3.2.

Table 3.2. Permutation V_{PW}^* values

Arrangement of Scores	Pair 1		Pair 2		Pair 3		Unpaired Observations			V_{PW}^*
	$g_S(X)$	$g_S(Y)$	$g_S(X)$	$g_S(Y)$	$g_S(X)$	$g_S(Y)$	$g_S(X)$	$g_S(X)$	$g_S(Y)$	
1	-0.143	0.543	0.086	0.543	-0.600	0.200	-0.800	-0.371	0.543	-1.828
2	-0.143	0.543	0.543	0.086	-0.600	0.200	-0.800	-0.371	0.543	-1.371
3	0.543	-0.143	0.086	0.543	-0.600	0.200	-0.800	-0.371	0.543	-1.142
4	-0.143	0.543	0.086	0.543	0.200	-0.600	-0.800	-0.371	0.543	-1.028
5	-0.143	0.543	0.086	0.543	-0.600	0.200	0.543	-0.800	-0.371	-0.914
6	0.543	-0.143	0.543	0.086	-0.600	0.200	-0.800	-0.371	0.543	-0.685
7	-0.143	0.543	0.543	0.086	0.200	-0.600	-0.800	-0.371	0.543	-0.571
8	-0.143	0.543	0.086	0.543	-0.600	0.200	0.543	-0.371	-0.800	-0.485
9	-0.143	0.543	0.543	0.086	-0.600	0.200	0.543	-0.800	-0.371	-0.457
10	0.543	-0.143	0.086	0.543	0.200	-0.600	-0.800	-0.371	0.543	-0.342
11	0.543	-0.143	0.086	0.543	-0.600	0.200	0.543	-0.800	-0.371	-0.228
12	-0.143	0.543	0.086	0.543	0.200	-0.600	0.543	-0.800	-0.371	-0.114
13	-0.143	0.543	0.543	0.086	-0.600	0.200	0.543	-0.371	-0.800	-0.028
14	0.543	-0.143	0.543	0.086	0.200	-0.600	-0.800	-0.371	0.543	0.115
15	0.543	-0.143	0.086	0.543	-0.600	0.200	0.543	-0.371	-0.800	0.201
16	0.543	-0.143	0.543	0.086	-0.600	0.200	0.543	-0.800	-0.371	0.229
17	-0.143	0.543	0.086	0.543	0.200	-0.600	0.543	-0.371	-0.800	0.315
18	-0.143	0.543	0.543	0.086	0.200	-0.600	0.543	-0.800	-0.371	0.343
19	0.543	-0.143	0.086	0.543	0.200	-0.600	0.543	-0.800	-0.371	0.572
20	0.543	-0.143	0.543	0.086	-0.600	0.200	0.543	-0.371	-0.800	0.658
21	-0.143	0.543	0.543	0.086	0.200	-0.600	0.543	-0.371	-0.800	0.772
22	0.543	-0.143	0.086	0.543	0.200	-0.600	0.543	-0.371	-0.800	1.001
23	0.543	-0.143	0.543	0.086	0.200	-0.600	0.543	-0.800	-0.371	1.029
24	0.543	-0.143	0.543	0.086	0.200	-0.600	0.543	-0.371	-0.800	1.458

From Table 3.2, there are six V_{PW}^* values that are greater than or equal to $V_{PW_{obs}}$, so

the p-value is $\frac{6}{24} = 0.250$.

Tables 3.3 and 3.4 illustrate the D - AK procedure.

Table 3.3. Computations for the V_{AK} statistic

X_i	δ_{x_i}	Y_i	δ_{y_i}	$\hat{S}_X(X_i)$	$\hat{S}_Y(X_i)$	$\bar{\hat{S}}(X_i)$	$h_{\bar{S}}(X_i)$	$\hat{S}_X(Y_i)$	$\hat{S}_Y(Y_i)$	$\bar{\hat{S}}(Y_i)$	$h_{\bar{S}}(Y_i)$
41	1	22	1	0.000	0.250	0.125	0.875	0.800	0.250	0.525	0.475
32	1	40	0	0.533	0.250	0.392	0.608	0.533	0.250	0.392	0.804
15	0	14	1	0.800	0.750	0.775	0.613	0.800	0.750	0.775	0.225
35	0	-	-	0.533	0.250	0.392	0.804	—	—	—	—
12	1	-	-	0.800	1.000	0.900	0.100	—	—	—	—
-	-	16	1	—	—	—	—	0.800	0.500	0.650	0.350

From Table 3.3, $V_{AK_{obs}} = \sum_{i=1}^5 h_{\bar{S}}(X_i) = 3.000$. The 24 equally likely V_{AK}^* values are given in Table 3.4.

Table 3.4. Permutation V_{AK}^* values

Arrangement of Scores	Pair 1		Pair 2		Pair 3		Unpaired Observations			V_{AK}^*
	$h_{\bar{S}}(X)$	$h_{\bar{S}}(Y)$	$h_{\bar{S}}(X)$	$h_{\bar{S}}(Y)$	$h_{\bar{S}}(X)$	$h_{\bar{S}}(Y)$	$h_{\bar{S}}(X)$	$h_{\bar{S}}(X)$	$h_{\bar{S}}(Y)$	
1	0.475	0.875	0.608	0.804	0.225	0.613	0.100	0.350	0.804	1.758
2	0.475	0.875	0.804	0.608	0.225	0.613	0.100	0.350	0.804	1.964
3	0.475	0.875	0.608	0.804	0.613	0.225	0.100	0.350	0.804	2.146
4	0.875	0.475	0.608	0.804	0.225	0.613	0.100	0.350	0.804	2.158
5	0.475	0.875	0.608	0.804	0.225	0.613	0.804	0.100	0.350	2.212
6	0.475	0.875	0.804	0.608	0.613	0.225	0.100	0.350	0.804	2.342
7	0.875	0.475	0.804	0.608	0.225	0.613	0.100	0.350	0.804	2.354
8	0.475	0.875	0.804	0.608	0.225	0.613	0.804	0.100	0.350	2.408
9	0.475	0.875	0.608	0.804	0.225	0.613	0.804	0.350	0.100	2.462
10	0.875	0.475	0.608	0.804	0.613	0.225	0.100	0.350	0.804	2.546
11	0.475	0.875	0.608	0.804	0.613	0.225	0.804	0.100	0.350	2.600
12	0.875	0.475	0.608	0.804	0.225	0.613	0.804	0.100	0.350	2.612
13	0.475	0.875	0.804	0.608	0.225	0.613	0.804	0.350	0.100	2.658
14	0.875	0.475	0.804	0.608	0.613	0.225	0.100	0.350	0.804	2.742
15	0.475	0.875	0.804	0.608	0.613	0.225	0.804	0.100	0.350	2.796
16	0.875	0.475	0.804	0.608	0.225	0.613	0.804	0.100	0.350	2.808
17	0.475	0.875	0.608	0.804	0.613	0.225	0.804	0.350	0.100	2.850
18	0.875	0.475	0.608	0.804	0.225	0.613	0.804	0.350	0.100	2.862
19	0.875	0.475	0.608	0.804	0.613	0.225	0.804	0.100	0.350	3.000
20	0.475	0.875	0.804	0.608	0.613	0.225	0.804	0.350	0.100	3.046
21	0.875	0.475	0.804	0.608	0.225	0.613	0.804	0.350	0.100	3.058
22	0.875	0.475	0.804	0.608	0.613	0.225	0.804	0.100	0.350	3.196
23	0.875	0.475	0.608	0.804	0.613	0.225	0.804	0.350	0.100	3.250
24	0.875	0.475	0.804	0.608	0.613	0.225	0.804	0.350	0.100	3.446

From Table 3.4, there are six V_{AK}^* values that are greater than or equal to $V_{AK_{obs}}$, so the p-value is $\frac{6}{24} = 0.250$.

Obviously, the computations for the $D-PW$ and $D-AK$ testing procedures will become quite laborious as the sample sizes become larger. Efficient computing routines are required for practical implementation of the tests. In light of this fact, we developed Fortran programs to execute the tests. These programs can be found in Appendix B.

3.7 Overview of the $D-PW$ and $D-AK$ Procedures

In this chapter we have developed $D-PW$ and $D-AK$, two different tests of H_0 that incorporate both paired and unpaired right censored observations. $D-PW$ and $D-AK$ each utilize a particular scoring system based on pooled ranks to calculate the associated test statistic. Both testing procedures utilize the permutation distribution of the test statistic based on a novel method of permuting the scores. Thus, both proposed tests possess all of the advantages and properties of permutation tests which are discussed in Section 3.2. The tests can be implemented using the computer routine described in Appendix B. Permutation versions of tests for paired right censored observations and for two independent right censored samples which use the proposed scoring systems are obtained as special cases of our test procedures.

CHAPTER 4 SIMULATION STUDY

4.1 Introduction

Closed form expressions for the power functions of the *D-PW* and *D-AK* tests do not exist. Therefore, finite sample size measures of performance of the *D-PW* and *D-AK* tests were investigated in a simulation study. Section 4.2 describes the general framework for the simulation study. The results are given in Section 4.3.

4.2 Description

Monte Carlo simulations were performed to investigate the powers of the *D-PW* and *D-AK* tests. The method used in simulating the situations studied are largely based on the methods used by O'Brien and Fleming (1987), Akritas (1992), and Woolson and O'Gorman (1992).

Survival times were generated under the model:

$$\begin{aligned}X'_i &= Z_i + Z_{2n+i}, \\Y'_i &= Z_{n+i} + Z_{2n+i}, \\X'_{n+j} &= Z_{4n+j} + Z_{4n+s+t+j}, \\Y'_{n+k} &= Z_{4n+s+k} + Z_{4n+2s+t+k},\end{aligned}$$

$i = 1, 2, \dots, n$; $j = 1, 2, \dots, s$; $k = 1, 2, \dots, t$. Thus, the Z_{2n+i} terms provide correlation among the paired observations. Censoring times were generated under the model:

$$C_i = cZ_{3n+i},$$

$$C_{n+j} = cZ_{4n+2s+2t+j},$$

$$C_{n+s+k} = cZ_{4n+3s+2t+k},$$

$i = 1, 2, \dots, n$; $j = 1, 2, \dots, s$; $k = 1, 2, \dots, t$. The Z_{2n+i} , $Z_{4n+s+t+j}$, $Z_{4n+2s+t+k}$, Z_{3n+i} , $Z_{4n+2s+2t+j}$, and $Z_{4n+3s+2t+k}$ terms were independent identically distributed (i.i.d.) unit exponential variates. The Z_i , Z_{n+i} , Z_{4n+j} , and Z_{4n+s+k} varied in their distributions according to the situation studied. The constant c determines the heaviness of censoring and thus also varied.

We evaluated the performance of the D -PW and D -AK tests under two null distributions and five alternatives to these null distributions.

The following null distributions were used.

- The Z_i , Z_{n+i} , Z_{4n+j} , and Z_{4n+s+k} were generated as i.i.d. unit exponential variates. This case will be denoted by *Exp*.
- The Z_i , Z_{n+i} , Z_{4n+j} , and Z_{4n+s+k} were generated as i.i.d. log-logistic variates by generating i.i.d. uniform(0,1) variates, U_i , U_{n+i} , U_{4n+j} , and U_{4n+s+k} , and then utilizing $U^{-1} - 1$. This case will be denoted by *LL*.

The following distributions were used as alternatives to the null distributions.

- The Z_i , Z_{n+i} , Z_{4n+j} , and Z_{4n+s+k} were generated as i.i.d. unit exponential variates, then each Z_i and Z_{4n+j} was multiplied by 5. This shift of scale increases the skewness in favor of the X sample (see Figure 4.1). The scale shift is frequently of interest for survival distributions. This case will be denoted by *Exp Sc.*

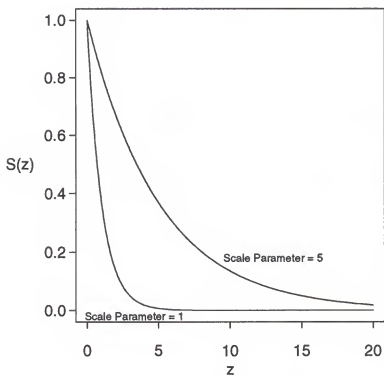


Figure 4.1. Exponential Scale Survival Functions

- The Z_i , Z_{n+i} , Z_{4n+j} , and Z_{4n+s+k} were generated as i.i.d. unit exponential variates, then 1 was added to each Z_i and Z_{4n+j} . This shifts location in favor of the X sample (see Figure 4.2). The location parameter often indicates an initial threshold or guarantee parameter before which it is assumed that failure cannot occur. This case will be denoted by *Exp Loc*.

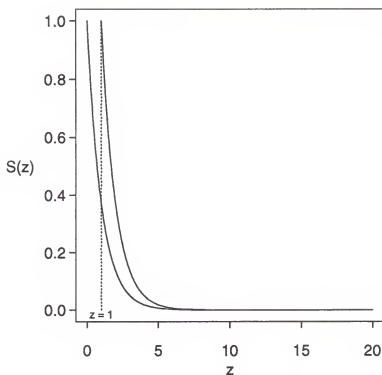


Figure 4.2. Exponential Location Survival Functions

- The Z_i , Z_{n+i} , Z_{4n+j} , and Z_{4n+s+k} were generated as i.i.d. unit exponential variates, then each Z_i and Z_{4n+j} was first multiplied by 1.5, and then raised to the power of 3. This shifts both the scale and the shape of the distribution in favor of the X sample (see Figure 4.3). The distribution of an exponential variate with scale parameter β raised to a positive power γ is generalized exponential, or Weibull with scale parameter β and shape parameter γ . This distribution is an important generalization of the exponential distribution that allows the hazard function to depend on time. This case will be denoted by *Gen Exp*.

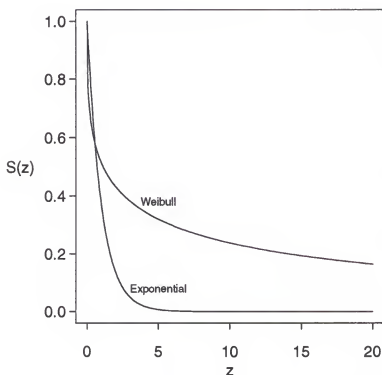


Figure 4.3. Generalized Exponential Survival Functions

- The Z_i , Z_{n+i} , Z_{4n+j} , and Z_{4n+s+k} were generated as i.i.d. log-logistic variates, then each Z_i and Z_{4n+j} was multiplied by 5. This shift of scale increases the skewness in favor of the X sample (see Figure 4.4). This case will be denoted by *LL Sc.*

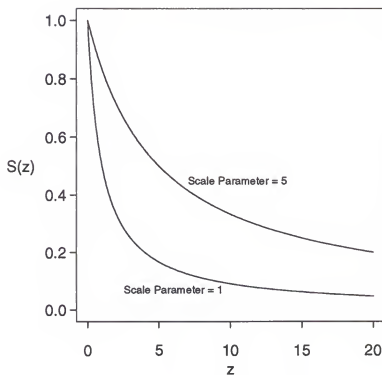


Figure 4.4. Log-logistic Scale Survival Functions

- The Z_i , Z_{n+i} , Z_{4n+j} , and Z_{4n+s+k} were generated as i.i.d. log-logistic variates, then 1 was added to each Z_i and Z_{4n+j} . This shifts location in favor of the X sample (see Figure 4.5). This case will be denoted by *LL Loc.*

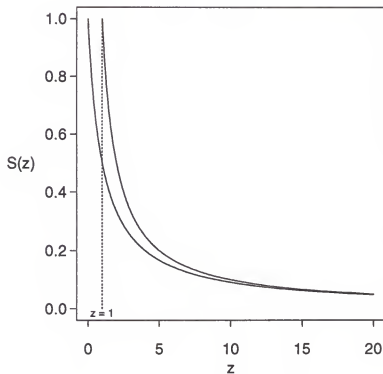


Figure 4.5. Log-logistic Location Survival Functions

For each distribution, performance was investigated under both 30% and 70% censoring, resulting in a total of 14 situations studied. To achieve the desired censoring percentage, the values of c were determined under the null distributions used; c was 5.1 for all *Exp* cases and 8.7 for all *LL* cases to give approximately 30% censoring, and c was 1.2 for all *Exp* cases and 1.6 for all *LL* cases to give approximately 70% censoring.

The theoretical null correlation between the paired observations in the *Exp* case is 0.5, as determined by the model described earlier. In several instances, we also considered correlations of 0.2 and 0.8 by appropriately scaling the distributions of Z_{2n+i} , $Z_{4n+s+t+j}$, and $Z_{4n+2s+t+k}$, and then adjusting c to maintain the desired censoring percentages. The theoretical correlation in the *LL* case is undefined, since the log-logistic distribution used has infinite variance.

We considered 33 different sample sizes according to $n \in \{5, 10, 25\}$ and $(s, t) \in \{(0, 5), (5, 0), (5, 5), (5, 10), (5, 20), (10, 5), (10, 10), (10, 20), (20, 5), (20, 10), (20, 20)\}$. Since randomly occurring unpaired observations may cause an extremely unbalanced (s, t) combination, such as $(5, 20)$, we felt it was important to investigate such combinations.

All results are based on simulations of 1000 samples from each sample size - situation combination. For each sample, the observed values of the test statistics, $V_{PW_{obs}}$ and $V_{AK_{obs}}$, and the permutation V_{PW}^* and V_{AK}^* values were computed. Whenever n , s , and t were such that $M \leq 5000$, each test utilized its entire permutation distribution; otherwise a random sample of size 5000 from its permutation distribution was utilized. Thus, the number of permutations used was $\tilde{M} = \min(M, 5000)$. The one-sided p-values, $\hat{p}_{PW} = \#\{V_{PW}^* \geq V_{PW_{obs}}\}/\tilde{M}$ and $\hat{p}_{AK} = \#\{V_{AK}^* \geq V_{AK_{obs}}\}/\tilde{M}$, were computed for each sample.

Empirical powers for the tests were obtained at a nominal α -level by computing $\hat{P}_{PW} = \#\{\hat{p}_{PW} \leq \alpha\}/1000$ and $\hat{P}_{AK} = \#\{\hat{p}_{AK} \leq \alpha\}/1000$. We examined $\alpha = .01, .05, .10$. Since the distributions are discrete, exact .01, .05, .10 α -level tests do not

always exist, in which case we examined α -levels as close as possible to .01, .05, and .10.

In order to compare the *D-PW* and *D-AK* tests under the alternative distributions, large sample 95% confidence intervals for each $(P_{PW} - P_{AK})$ were calculated using

$$(\hat{P}_{PW} - \hat{P}_{AK}) \pm 1.96\sqrt{\hat{Var}(\hat{P}_{PW} - \hat{P}_{AK})}, \quad (4.1)$$

where $\hat{Var}(\hat{P}_{PW} - \hat{P}_{AK}) = \hat{Var}(\hat{P}_{PW}) + \hat{Var}(\hat{P}_{AK}) - 2\hat{Cov}(\hat{P}_{PW}, \hat{P}_{AK})$.

4.3 Results

Tables C.1 - C.33 in Appendix C contain the empirical powers under each situation for each sample size considered. Figures 4.6 - 4.23 reflect the contents of Tables C.1 - C.33. The situations studied are labeled on the horizontal axis by the corresponding distribution abbreviation and percent censoring. For instance, *Exp30* signifies the previously described *Exp* case with 30% censoring.

Tables C.1 - C.33 and Figures 4.6 - 4.23 indicate that the *D-PW* and *D-AK* procedures generally have similar power. For each sample size, relative performance of the procedures was consistent across situations. There are some sample sizes for which one procedure performed slightly better than the other. For $s < t$, i.e. when the number of observations generated to have longer survival times is less than the number generated to have shorter survival times, the *D-AK* procedure generally performed slightly better than the *D-PW* procedure. For $s > t$, the *D-PW* procedure generally performed slightly better. Figures 4.6 - 4.23 illustrate this phenomenon.

The nominal α -levels were basically maintained. For small sample sizes, the test statistics have highly discrete distributions, resulting in tests that are necessarily conservative. Note that the powers under H_0 are much smaller than the nominal level for $n = 5, s = 0, t = 5$ and $n = 5, s = 5, t = 0$. This is particularly evident under 70% censoring, even for the sample sizes $n = 10, s = 0, t = 5$ and $n = 10, s = 5, t = 0$.

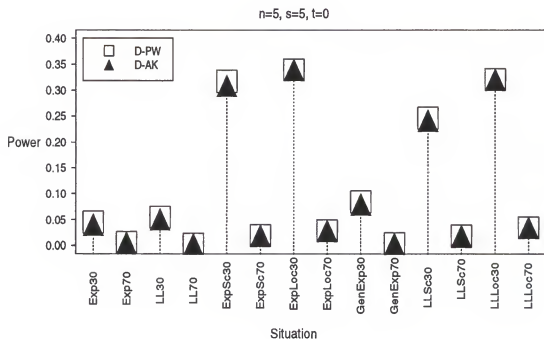
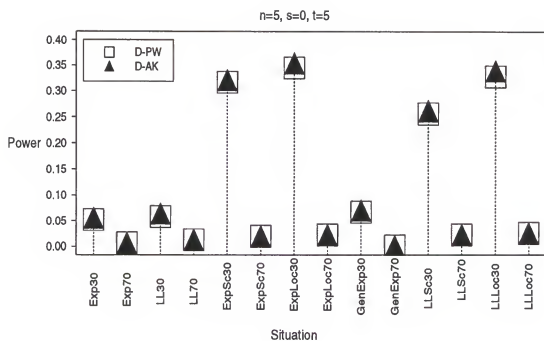


Figure 4.6. Empirical Powers for $n = 5, s = 0, t = 5$ and $n = 5, s = 5, t = 0$ at $\alpha = .093$

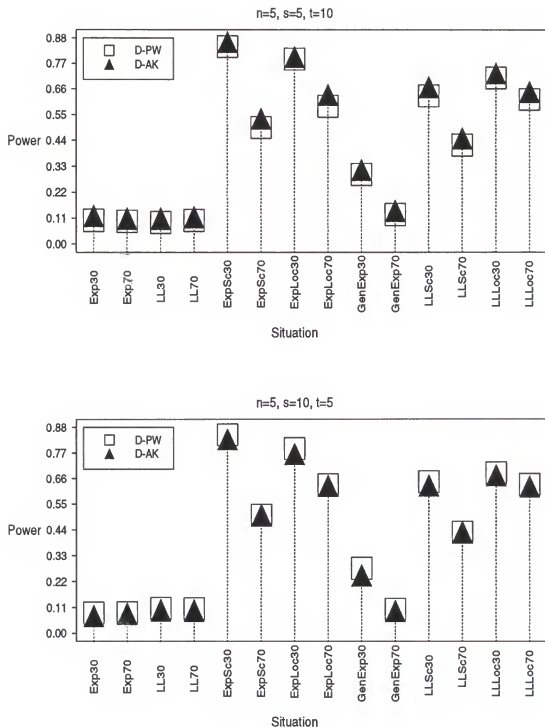


Figure 4.7. Empirical Powers for $n = 5, s = 5, t = 10$ and $n = 5, s = 10, t = 5$ at $\alpha = .10$

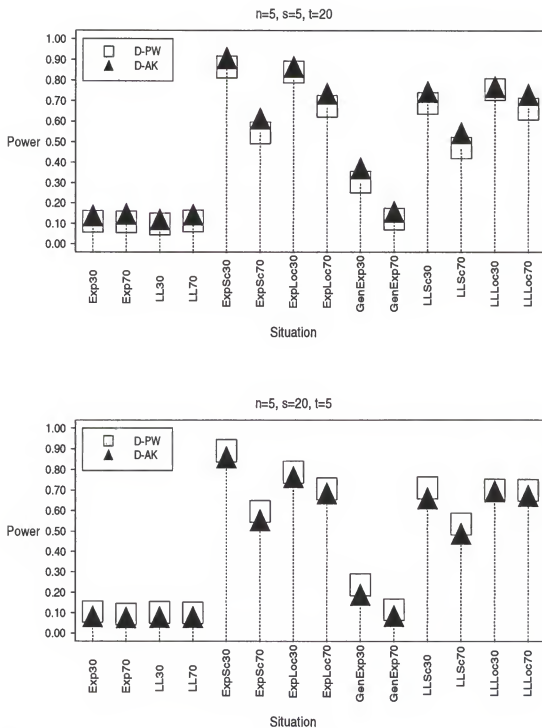


Figure 4.8. Empirical Powers for $n = 5, s = 5, t = 20$ and $n = 5, s = 20, t = 5$ at $\alpha = .10$

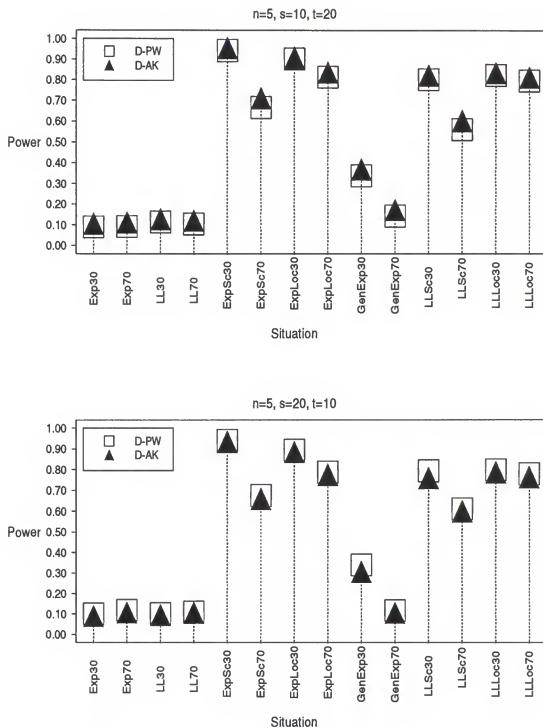


Figure 4.9. Empirical Powers for $n = 5$, $s = 10$, $t = 20$ and $n = 5$, $s = 20$, $t = 10$ at $\alpha = .10$

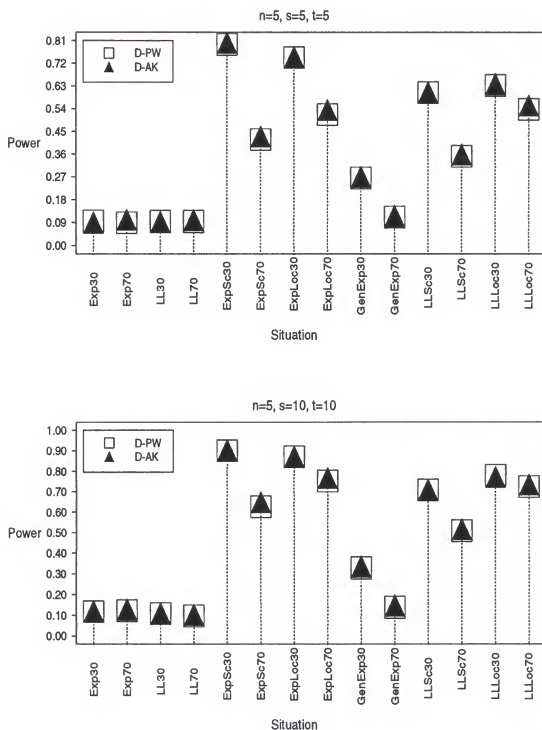


Figure 4.10. Empirical Powers for $n = 5, s = 5, t = 5$ and $n = 5, s = 10, t = 10$ at $\alpha = .10$

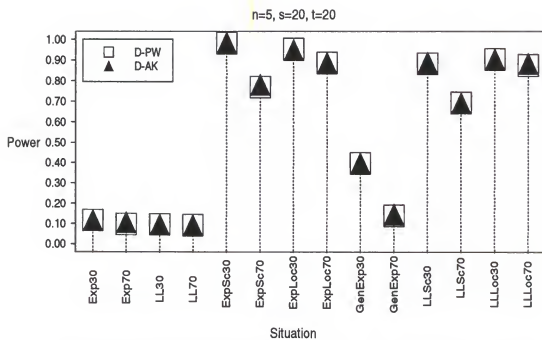


Figure 4.11. Empirical Powers for $n = 5, s = 20, t = 20$ at $\alpha = .10$

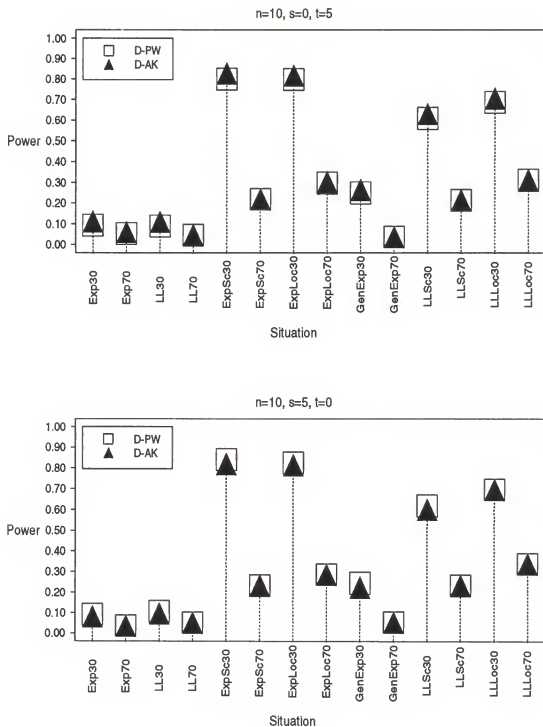


Figure 4.12. Empirical Powers for $n = 10, s = 0, t = 5$ and $n = 10, s = 5, t = 0$ at $\alpha = .101$

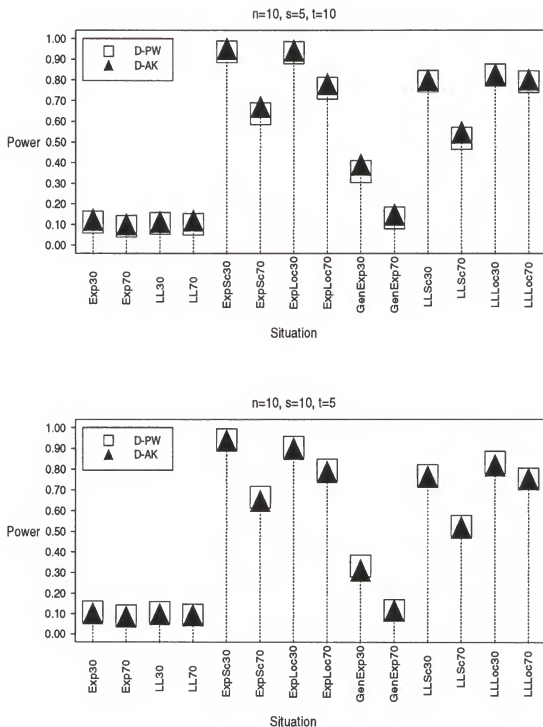


Figure 4.13. Empirical Powers for $n = 10, s = 5, t = 10$ and $n = 10, s = 10, t = 5$ at $\alpha = .10$

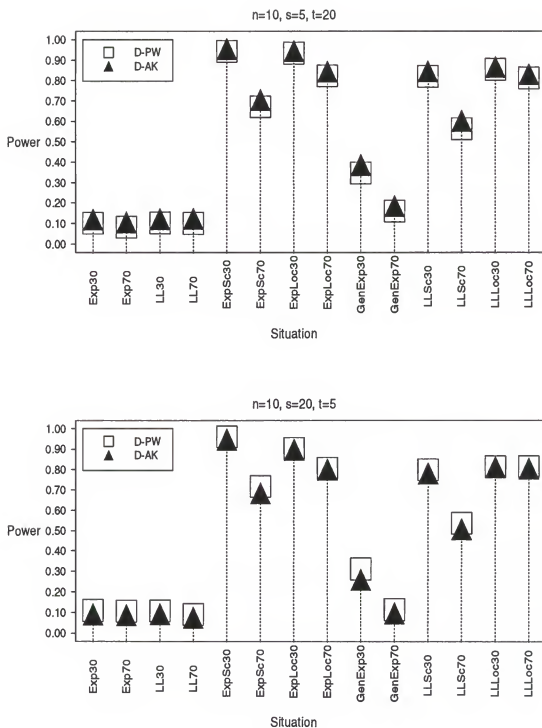


Figure 4.14. Empirical Powers for $n = 10, s = 5, t = 20$ and $n = 10, s = 20, t = 5$ at $\alpha = .10$

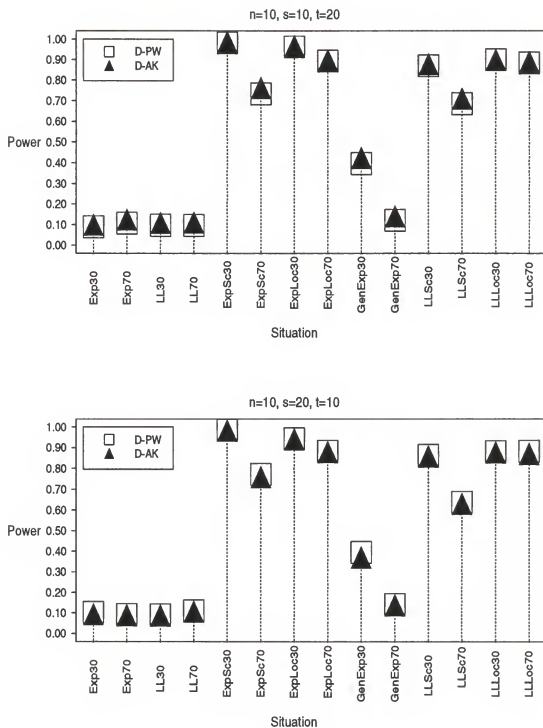


Figure 4.15. Empirical Powers for $n = 10, s = 10, t = 20$ and $n = 10, s = 20, t = 10$ at $\alpha = .10$

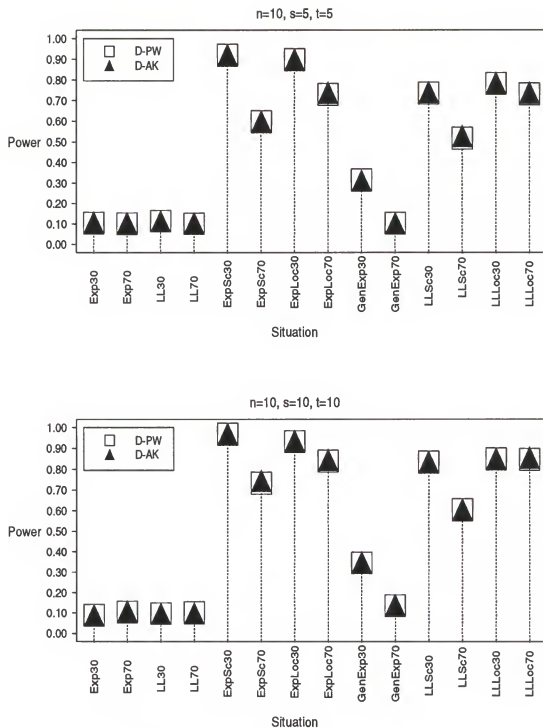


Figure 4.16. Empirical Powers for $n = 10, s = 5, t = 5$ and $n = 10, s = 10, t = 10$ at $\alpha = .10$

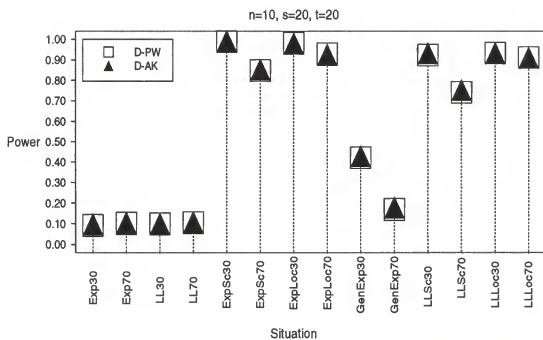


Figure 4.17. Empirical Powers for $n = 10, s = 20, t = 20$ at $\alpha = .10$

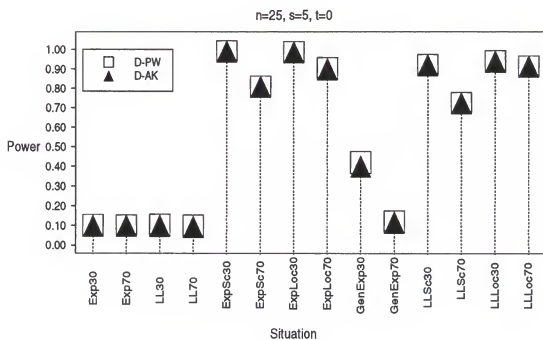
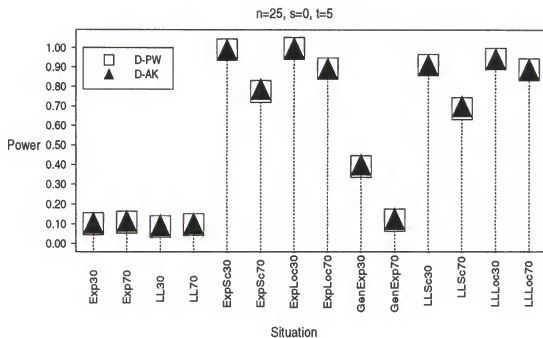


Figure 4.18. Empirical Powers for $n = 25, s = 0, t = 5$ and $n = 25, s = 5, t = 0$ at $\alpha = .10$

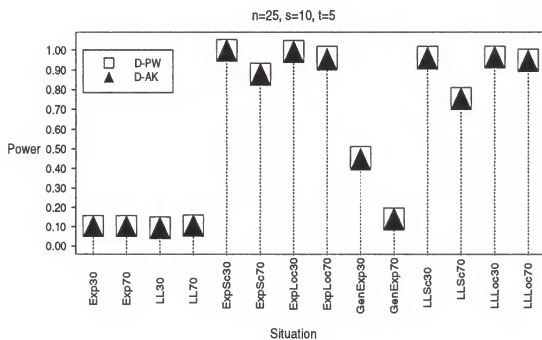
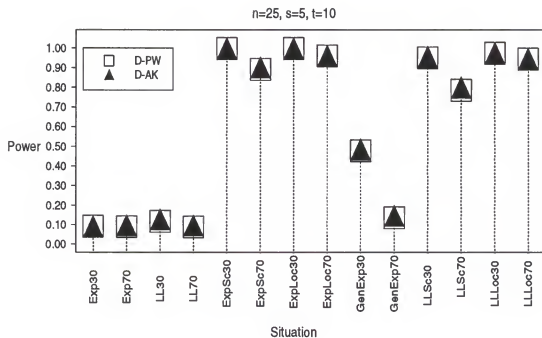


Figure 4.19. Empirical Powers for $n = 25, s = 5, t = 10$ and $n = 25, s = 10, t = 5$ at $\alpha = .10$

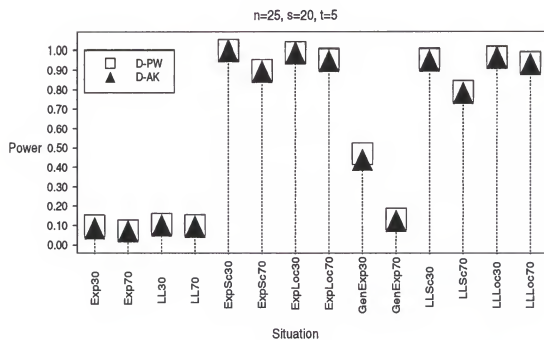
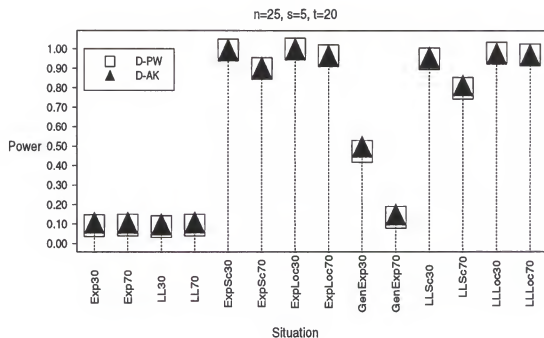


Figure 4.20. Empirical Powers for $n = 25, s = 5, t = 20$ and $n = 25, s = 20, t = 5$ at $\alpha = .10$

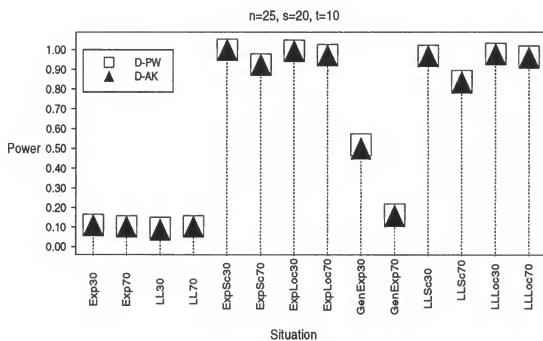
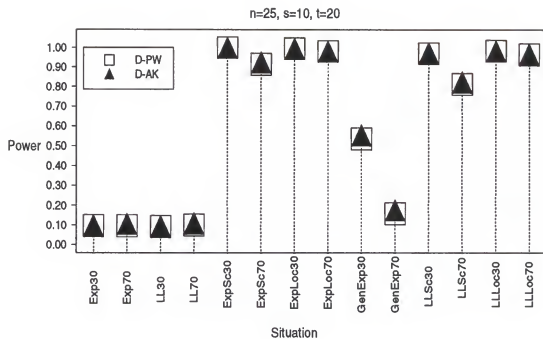


Figure 4.21. Empirical Powers for $n = 25, s = 10, t = 20$ and $n = 25, s = 20, t = 10$ at $\alpha = .10$

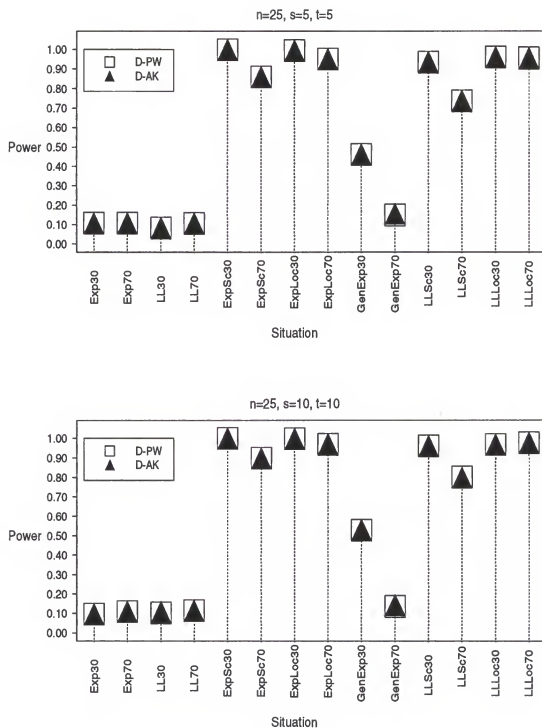


Figure 4.22. Empirical Powers for $n = 25, s = 5, t = 5$ and $n = 25, s = 10, t = 10$ at $\alpha = .10$

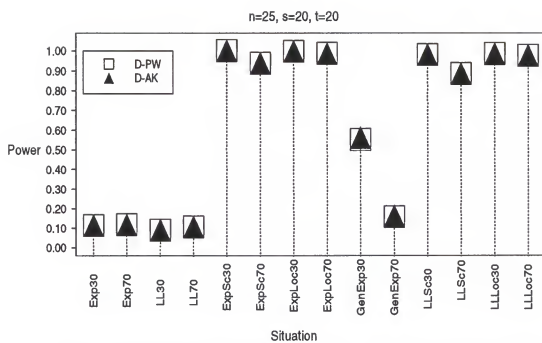


Figure 4.23. Empirical Powers for $n = 25, s = 20, t = 20$ at $\alpha = .10$

The *D-PW* and *D-AK* procedures demonstrated broad sensitivity for the alternatives studied. Over the range of sample sizes studied, the test procedures demonstrated the ability to detect scale and location shifts in both the exponential distribution and log-logistic distribution with high power, under both 30% and 70% censoring. The procedures detected the Weibull alternative to the exponential distribution with moderate power for large sample sizes under 30% censoring.

We found that as censoring increased from 30% to 70%, the power decreased for each alternative distribution. This was especially noticeable for the Weibull (Generalized Exponential) alternative.

The following table helps to illustrate the above mentioned findings. Because of the general similarity of the *D-PW* and *D-AK* empirical powers, we collapsed these powers across all sample sizes studied into a single sample for each situation, and tabled the 0th, 25th, 50th, 75th, and 100th percentiles. The tabled entries were obtained at a nominal $\alpha = (\approx) 0.10$. The general patterns depicted by Figures 4.6 - 4.23 and Table 4.1 also hold under correlations of $\rho = 0.2$ and $\rho = 0.8$. As expected, we observed that power is inversely related to the degree of correlation. For example, at $n = 25$, $s = 5$, and $t = 10$, under *Exp Sc*, 70, and at $\alpha = .01$, the empirical power of the *D-PW* [*D-AK*] test is .543 [.562] at $\rho = 0.2$, .514 [.536] at $\rho = 0.5$, and .489 [.508] at $\rho = 0.8$.

Table 4.1 Percentiles of Empirical Power Estimates of both the *D-PW* and *D-AK* Tests, Across all 33 Sample Sizes, at $\alpha =$ (or \approx) .10.

	Situation	Percentiles				
		0	25	50	75	100
H_0 True	<i>Exp, 30</i>	.0390	.0910	.1005	.1090	.1400
	<i>Exp, 70</i>	.0050	.0880	.1010	.1070	.1460
	<i>LL, 30</i>	.0500	.0920	.0975	.1060	.1270
	<i>LL, 70</i>	.0020	.0940	.1010	.1080	.1440
H_a True	<i>Exp Sc, 30</i>	.3080	.8900	.9580	.9950	1.000
	<i>Exp Sc, 70</i>	.0190	.5960	.7140	.8620	.9400
	<i>Exp Loc, 30</i>	.3390	.8410	.9340	.9940	.9990
	<i>Exp Loc, 70</i>	.0220	.7060	.8375	.9500	.9890
	<i>Gen Exp, 30</i>	.0670	.3040	.3695	.4470	.5620
	<i>Gen Exp, 70</i>	.0010	.1140	.1345	.1470	.1840
	<i>LL Sc, 30</i>	.2410	.7070	.8270	.9350	.9850
	<i>LL Sc, 70</i>	.0180	.5060	.6025	.7510	.8870
	<i>LL Loc, 30</i>	.3200	.7570	.8495	.9620	.9870
	<i>LL Loc, 70</i>	.0260	.6980	.8220	.9370	.9780

4.4 Explanation of Results

It is not a total surprise that the *D-PW* and *D-AK* procedures are powerful for detecting scale shifts, since the scores we propose are related to the optimal scores for detecting scale shifts in the log-logistic distribution in the two-sample case. The simulation study indicates that desirable power properties for scale shifts are maintained when there exists correlation between the samples via the paired observations. In addition to scale shifts, location shifts are generally well-detected. A similar phenomenon was observed in the simulation studies of O'Brien and Fleming (1987) and Woolson and O'Gorman (1992), who used scores similar to ours in pooled rank tests for paired data. The Weibull alternative can also be detected, but with moderate power provided large sample sizes and a low amount of censoring. This is most likely due to the fact that in the Weibull case, when the shape parameter differs the survival functions cross, a situation where it is known that the Prentice-Wilcoxon (and other linear rank) scores do not produce powerful tests (Lee, 1992; Sun and Sherman, 1996).

Standard rank test procedures that accomodate censored observations generally experience a decrease in power for detecting alternatives as censoring increases. This is due to the fact that each censored observation between two adjacent uncensored observations is assigned the same score. Thus, as censoring increases, the number of distinct values of the test statistic decreases, resulting in a decrease in power. Consistent with this fact, we found that as censoring increased, power for detecting alternatives decreased.

CHAPTER 5 APPLICATION AND DISCUSSION

5.1 Introduction

The *D-PW* and *D-AK* tests are illustrated with a real data set in Section 5.2. Section 5.3 illustrates advantages of using the *D-PW* and *D-AK* tests over corresponding tests which use only the paired observations. Section 5.4 presents possible extensions to the research. Recommendations for use of the tests and conclusions of the dissertation are given in Section 5.5.

5.2 Analysis of Skin Graft Data

Batchelor and Hackett (1970) reported the survival times of closely and poorly matched human lymphocyte antigen (HLA) matched skin grafts on the same burned individual. A slightly modified form of the data set follows.

For this data set, $n = 11$, $s = 1$, and $t = 4$; so $M = 2^{11} \times \binom{1+4}{1} = 10,240$.

Calculating the respective scores, it follows that $V_{PW_{obs}} = 2.67857$, while $V_{AK_{obs}} = 7.53889$. Utilizing the program described in the Appendix, the exact p-values based on all 10,240 permutations were computed. The (one-sided) p-value for the *D-PW* procedure is 0.019141, while the (one-sided) p-value for the *D-AK* procedure is 0.017578. The p-values suggest strong evidence that closely matched grafts survive longer than poorly matched grafts.

Table 5.1 Days of Survival. Censored observations are indicated by + signs.

Patient	Close Match (<i>X</i> Sample)	Poor Match (<i>Y</i> Sample)
1	—	19
2	24	—
3	—	18
4	37	29
5	19	13
6	—	19
7	57 ⁺	15
8	93	26
9	16	11
10	22	17
11	20	26
12	18	21
13	63	43
14	—	28 ⁺
15	29	15
16	60 ⁺	40

Technical difficulties in the experiment resulted in unpaired observations for Patients 1, 2, 3, 6, and 14.

5.3 The Effect of Discarding Unpaired Observations

A subset of the data set listed in Table 5.1 has been analyzed by many researchers who discarded the randomly occurring unpaired observations and utilized test procedures applicable to paired right censored data. Following along these lines, using only the paired data from Table 5.1, we applied the paired data (asymptotic) test procedures of O'Brien and Fleming and Akritas, as described in Section 2.3. These procedures are based on scoring systems corresponding to the *D-PW* and *D-AK* procedures respectively, and will be called the Paired *PW* Test and the Paired *AK* Test respectively. For the Paired *PW* Test we obtained a p -value = 0.018993, while for the Paired *AK* Test we obtained a p -value = 0.014184. Since the p -values from these Paired Tests only slightly differ from the p -values from the corresponding *D-PW* and *D-AK* Tests, discarding the randomly occurring unpaired observations has

little effect on the p-value for the given data set. However, this will not be true in general.

To investigate the effect of discarding randomly occurring unpaired observations, we utilized the Monte Carlo simulation framework described in Section 4.2 with $n \in \{5, 10, 25\}$ and $(s, t) = (0, 0)$. We again used 1000 samples from each sample size - situation combination. For each sample, the one-sided p-values for the Paired *PW* Test and the Paired *AK* Test were computed. Empirical powers were obtained under a null correlation of 0.5, at a nominal .10 level, in the fashion described in Section 4.2.

The powers of the Paired *PW* and *AK* Tests are given in Table D.1 in Appendix D. These powers were compared to the corresponding powers of the *D-PW* and *D-AK* procedures at $\alpha = .10$ (or $\approx .10$) with $(s, t) \in \{(0, 5), (5, 0), (5, 5), (5, 10), (5, 20), (10, 5), (10, 10), (10, 20), (20, 5), (20, 10), (20, 20)\}$, which are given in Tables C.1 - C.33 in Appendix C. The results of these comparisons are graphically displayed by Figures 5.1 - 5.6.

In Figures 5.1 - 5.6, points to the right of the vertical line represent power differences under the alternative hypotheses. Therefore, right of the vertical line, points that lie above 0 represent sample sizes for which the *D-PW* and *D-AK* tests have greater power than the corresponding tests that discard unpaired observations. The figures indicate that the powers of the *D-PW* and *D-AK* tests are greater than the corresponding paired tests with very few exceptions. These exceptions primarily represent the (s, t) combinations of $(0, 5)$ or $(5, 0)$. The reason for the negative power difference with these (s, t) combinations is the unavoidable conservativeness of exact tests with small sample sizes; these (s, t) combinations do not contribute to the number of permutations, M , which simplifies to 2^n in these cases. An example of this conservativeness is seen at $n = 5, s = 5, t = 0$, under *Exp, 70*, where the empirical power of the *D-PW* test is .005, considerably lower than its nominal .093 level. In contrast, the corresponding empirical power of the Paired *PW* test is .110 at a nominal .10 level.

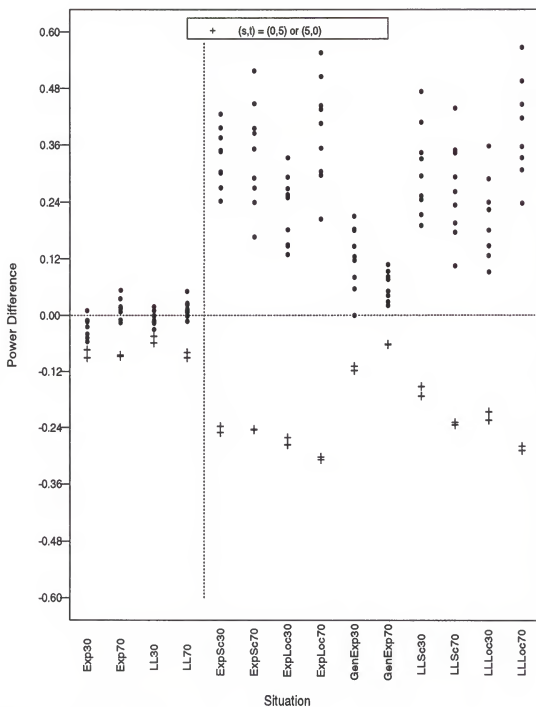


Figure 5.2. Power Differences ($D-AK$ Minus Paired AK Test) with $n = 5$ at $\alpha = .10$ (or $\approx .10$)

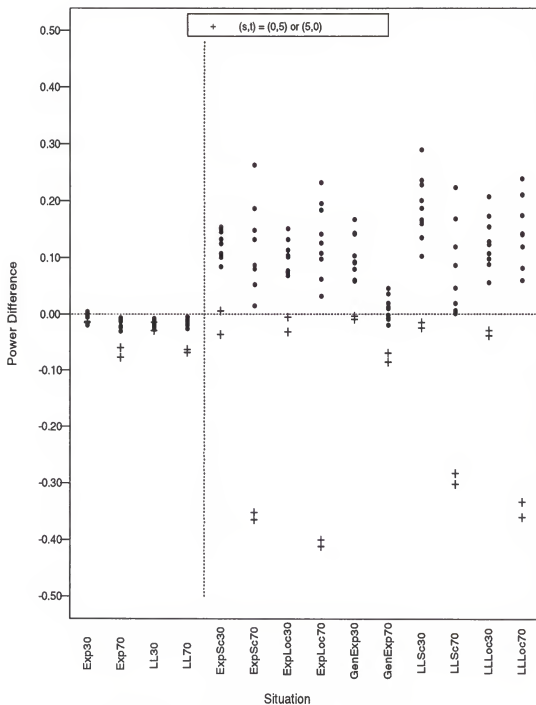


Figure 5.3. Power Differences ($D-PW$ Minus Paired PW Test) with $n = 10$ at $\alpha = .10$ (or $\approx .10$)

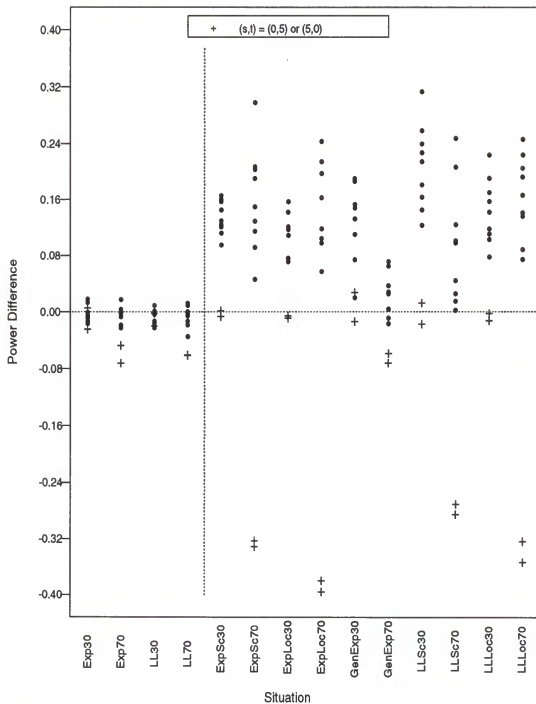


Figure 5.4. Power Differences ($D-AK$ Minus Paired AK Test) with $n = 10$ at $\alpha = .10$ (or $\approx .10$)

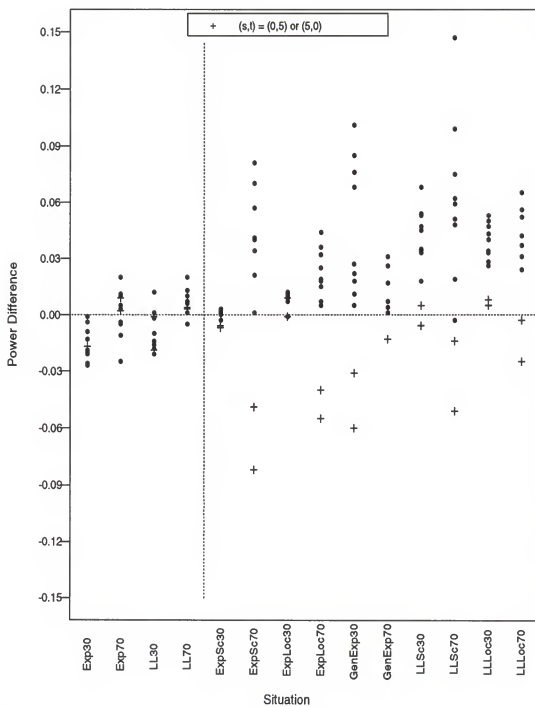


Figure 5.5. Power Differences (D -PW Minus Paired PW Test) with $n = 25$ at $\alpha = .10$

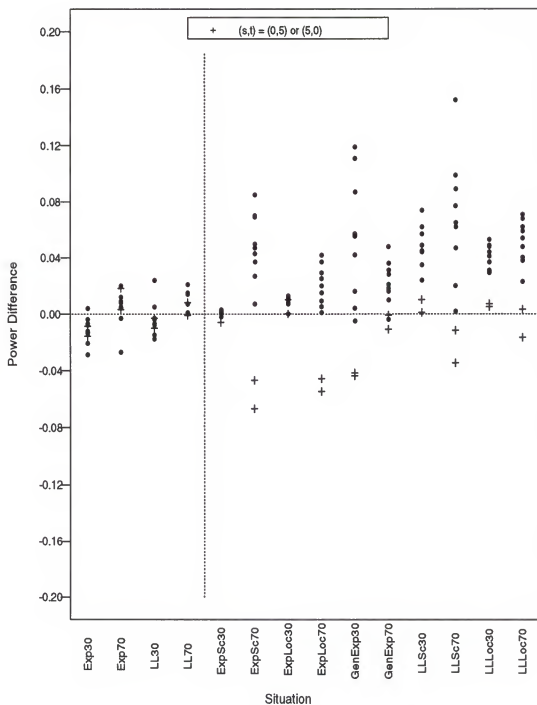


Figure 5.6. Power Differences ($D-AK$ Minus Paired AK Test) with $n = 25$ at $\alpha = .10$

As expected, the power difference increases as s and t increase at fixed n . Also, the figures show that as n increases, the unpaired observations have less impact on the difference in power. The comparisons were made assuming a null correlation of $\rho = 0.5$ for the paired observations. As ρ increases, the unpaired observations become more crucial for detecting alternatives. Thus, the larger the value of ρ , we can expect a greater the power difference between D -PW [D -AK] and the corresponding paired test, in favor of the D -PW [D -AK] test.

In conclusion, the power difference is a function of n, s, t , and ρ . Discarding randomly occurring unpaired observations and applying tests for paired data can have a detrimental effect on power when compared to the D -PW and D -AK tests, especially for small values of n , and large values of s, t , and ρ .

5.4 Extensions to the Research

There are other avenues to be investigated. For instance, it might be desirable to investigate the performance of V when scoring systems other than those proposed are used. More generally, because we have employed pooled rank procedures, any two-sample statistic that accomodates right censored data is a candidate test statistic.

Specifically, it may be interesting to examine the performance of tests based on the statistic

$$W = \max_k \sup_t \left| \sum_{t_j \leq t} w_{j,k} \left(\frac{d_{j,k}}{r_{j,k}} - \frac{d_j}{r_j} \right) \right|,$$

where k indexes the sample, t_j denotes the uncensored survival times in the combined sample, $w_{j,k}$ is a weight function that can be chosen to emphasize early or late deaths, $d_{j,k}$ denotes the number of deaths at time t_j for sample k , $r_{j,k}$ denotes the number in the risk set at time t_j for sample k , d_j denotes the number of deaths at time t_j in the combined sample, and r_j denotes the number in the risk set at time t_j in the combined sample. Sun and Sherman (1996) noted that W is sensitive to differences that occur between the two samples at any time t . These authors showed that permutation

tests based on W for the two-sample problem are competitive with the standard two-sample nonparametric tests and superior under certain alternatives. To use the test based on W for the type of data we consider, that is, data consisting of both paired and unpaired observations, one would permute the data in the fashion described in Section 3.6 in order to calculate W^* values.

It may also be interesting to investigate conditions under which the permutation method would be appropriate when allowing the censoring time to differ for members of the same pair. This censoring mechanism is more realistic in many studies and less restrictive than the conventional common censoring time assumption.

5.5 Recommendations and Conclusions

There exist no established tests for the equality of two survival distributions on the basis of randomly right censored data consisting of both paired and unpaired observations. In this dissertation, we have developed two such test procedures, provided computer programs that are essential for easy implementation of the procedures, and demonstrated the usefulness of the procedures.

In general, the tests are recommended for detecting scale and location shifts. For Weibull alternatives, the tests are recommended for use only with large sample sizes and low amounts of censoring. Based on our power study, the $D-AK$ test procedure is recommended when the number of survival times hypothesized to be longer is less than or equal to the number hypothesized to be shorter. Otherwise, the $D-PW$ test procedure is recommended. The tests should be used cautiously with small sample sizes, especially with a high amount of censoring, since they will be conservative. In many cases, our tests provide improved power over corresponding (asymptotic) paired data test procedures that discard randomly occurring unpaired observations.

APPENDIX A ASYMPTOTIC DISTRIBUTIONS OF V_{PW} AND V_{AK}

Following is the derivation of the asymptotic distributions of the V_{PW} and V_{AK} statistics under the null hypothesis. For this derivation, we assume that $(X_i, \delta_{xi}, Y_i, \delta_{yi})$, $(X_{n+j}, \delta_{x_{n+j}})$, $(Y_{n+k}, \delta_{y_{n+k}})$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, s$; $k = 1, 2, \dots, t$, are as in Section 2.2 and use the following notation:

$$\tilde{N} = 2n + s + t,$$

$$\alpha_j = I(\text{subject } j \text{ belongs to } X \text{ sample}),$$

$$B_{1p}(v) = \sum_{j=1}^n I(X_j > v), \quad B_{1u}(v) = \sum_{j=n+1}^{n+s} I(X_j > v),$$

$$B_{2p}(v) = \sum_{j=1}^n I(Y_j > v), \quad B_{2u}(v) = \sum_{j=n+1}^{n+t} I(Y_j > v),$$

$$B_1(v) = B_{1p}(v) + B_{1u}(v), \quad B_2(v) = B_{2p}(v) + B_{2u}(v),$$

$$B(v) = B_1(v) + B_2(v),$$

$$K(v) = W(v) \left\{ \frac{\tilde{N}}{(n+s)(n+t)} \right\}^{\frac{1}{2}} \frac{B_1(v)B_2(v)}{B(v)}, \quad W \text{ being an appropriately chosen weight function,}$$

$$H_{j1}(v) = \frac{K(v)}{B_1(v)} \alpha_j, \quad H_{j2}(v) = \frac{K(v)}{B_2(v)} (1 - \alpha_j),$$

$$N_j(v) = I(Z_j < v, \delta_{z_j} = 1), \quad N(v) = \sum_{j=1}^{\tilde{N}} N_j(v),$$

$$\lambda(u) = \frac{f(u)}{S(u)}, \quad \text{where } f, S \text{ are the density and survival functions respectively,}$$

$$\Lambda(v) = \int_0^v \lambda(u) du,$$

$$M_j(v) = N_j(v) - \int_0^v I(Z_j > u) d\Lambda(u),$$

$$\pi_{11}(v) = P(X > v, Y > v), \quad \pi_{10}(v) = P(X > v, Y < v), \quad \pi_{01}(v) = P(X < v, Y > v),$$

$$\pi_1(v) = P(X > v), \quad \pi_2(v) = P(Y > v).$$

A counting process representation, as in Fleming and Harrington (1991), is useful for finding the asymptotic distributions of V_{PW} and V_{AK} . These statistics can be expressed as

$$V' = U_1 - U_2,$$

where,

$$U_i \equiv U_i(\infty) = \sum_{j=1}^{\tilde{N}} \int_0^{\infty} H_{ji}(v) dM_j(v),$$

$i = 1, 2$. By choosing $W(v) = \tilde{S}(v)$, the pooled Kaplan-Meier estimator, we obtain V'_{PW} (see Fleming and Harrington). By choosing $W(v) = \bar{\tilde{S}}(v)$, the average Kaplan-Meier estimator, we obtain V'_{AK} .

Theorem 5.3.5 in Fleming and Harrington establishes sufficient conditions for a dependent counting process to converge to a multivariate Gaussian process. In our problem, due to the paired observations, the U_i are dependent counting processes. Thus, invoking Theorem 5.3.5, we obtain the following result pertaining to the asymptotic bivariate normality of U_1 and U_2 . The asymptotic null distributions of V'_{PW} and V_{AK} are stated in Result A.3.

Result A.1 Let $\tilde{N} \rightarrow \infty$ such that $\frac{2n}{\tilde{N}} \rightarrow a_1$, $\frac{s}{\tilde{N}} \rightarrow a_2$, $\frac{t}{\tilde{N}} \rightarrow a_3$, where $a_k \in (0, 1)$ and $\sum_{k=1}^3 a_k = 1$. Suppose that $W(v)$ converges in probability to $W_*(v)$. Then, under H_0 , for all $t^* > 0$,

$$(U_1(t^*), U_2(t^*)) \xrightarrow{D} N(\mu, \Sigma),$$

as $\tilde{N} \rightarrow \infty$, where $\mu = (0, 0)$, $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$, and

$$\sigma_{11} = \left(\frac{a_1 + 2a_3}{2} \right) \int_0^{t^*} W_*^2(v) \tilde{\pi}_1(v) d\Lambda(v),$$

$$\sigma_{22} = \left(\frac{a_1 + 2a_2}{2} \right) \int_0^{t^*} W_*^2(v) \tilde{\pi}_2(v) d\Lambda(v),$$

$$\sigma_{12} = 0.$$

Here,

$$\begin{aligned}\tilde{\pi}_1 &= \frac{[(\frac{a_1}{a_1+2a_2})(\pi_{11} + \pi_{10}) + (\frac{2a_2}{a_1+2a_2})\pi_1][(\frac{a_1}{a_1+2a_3})(\pi_{11} + \pi_{01}) + (\frac{2a_3}{a_1+2a_3})\pi_2]^2}{[a_1\pi_{11} + \frac{a_1}{2}\pi_{10} + \frac{a_1}{2}\pi_{01} + a_2\pi_1 + a_3\pi_2]^2} \\ \tilde{\pi}_2 &= \frac{[(\frac{a_1}{a_1+2a_2})(\pi_{11} + \pi_{10}) + (\frac{2a_2}{a_1+2a_2})\pi_1]^2[(\frac{a_1}{a_1+2a_3})(\pi_{11} + \pi_{01}) + (\frac{2a_3}{a_1+2a_3})\pi_2]}{[a_1\pi_{11} + \frac{a_1}{2}\pi_{10} + \frac{a_1}{2}\pi_{01} + a_2\pi_1 + a_3\pi_2]^2}\end{aligned}$$

Proof. Let,

$$Q_{ii} = \sum_{j=1}^{\tilde{N}} \int_0^{t^*} H_{ji}^2(v) I(Z_j > v) d\Lambda(v), \quad i = 1, 2,$$

$$Q_{12} = \sum_{j=1}^{\tilde{N}} \int_0^{t^*} H_{j1}(v) H_{j2}(v) I(Z_j > v) d\Lambda(v),$$

and

$$Q_{ii,\epsilon} = \sum_{j=1}^{\tilde{N}} \int_0^{t^*} H_{ji}^2(v) I(|H_{ji}(v)| > \epsilon) I(Z_j > v) d\Lambda(v), \quad i = 1, 2.$$

Using Theorem 5.3.5 in Fleming and Harrington, the proof is complete if we show

$$Q_{ii} \xrightarrow{P} \sigma_{ii}, \quad Q_{12} \xrightarrow{P} 0, \quad \text{and} \quad Q_{ii,\epsilon} \xrightarrow{P} 0.$$

Now, $Q_{12} = 0$, by definition of H_{ji} . Thus, $Q_{12} \xrightarrow{P} 0$. Also,

$$\begin{aligned}Q_{11} &= \sum_{j=1}^{\tilde{N}} \int_0^{t^*} H_{j1}^2(v) I(Z_j > v) d\Lambda(v) \\ &= \sum_{j=1}^{n+s} \int_0^{t^*} W^2(v) \frac{\tilde{N}}{(n+s)(n+t)} \left(\frac{B_{2p}(v) + B_{2u}(v)}{B_{1p}(v) + B_{1u}(v) + B_{2p}(v) + B_{2u}(v)} \right)^2 \times \\ &\quad I(X_j > v) d\Lambda(v) \\ &= \int_0^{t^*} W^2(v) \frac{\tilde{N}}{(n+s)(n+t)} \left(\frac{B_{2p}(v) + B_{2u}(v)}{B_{1p}(v) + B_{1u}(v) + B_{2p}(v) + B_{2u}(v)} \right)^2 \times \\ &\quad \{B_{1p}(v) + B_{1u}(v)\} d\Lambda(v)\end{aligned}$$

$$\begin{aligned}
&= \int_0^{t^*} \left(\frac{n+t}{\tilde{N}} \right) W^2(v) \left(\frac{\tilde{N}}{B_{1p}(v) + B_{1u}(v) + B_{2p}(v) + B_{2u}(v)} \right)^2 \times \\
&\quad \left(\frac{B_{1p}(v) + B_{1u}(v)}{n+s} \right) \left(\frac{B_{2p}(v) + B_{2u}(v)}{n+t} \right)^2 d\Lambda(v) \\
&\xrightarrow{P} \left(\frac{a_1 + 2a_3}{2} \right) \int_0^{t^*} W_*^2(v) \times \\
&\quad \frac{\left[\left(\frac{a_1}{a_1+2a_2} \right) (\pi_{11} + \pi_{10}) + \left(\frac{2a_2}{a_1+2a_2} \right) \pi_1 \right] \left[\left(\frac{a_1}{a_1+2a_3} \right) (\pi_{11} + \pi_{01}) + \left(\frac{2a_3}{a_1+2a_3} \right) \pi_2 \right]^2}{[a_1\pi_{11} + \frac{a_1}{2}\pi_{10} + \frac{a_1}{2}\pi_{01} + a_2\pi_1 + a_3\pi_2]^2} d\Lambda(v),
\end{aligned}$$

since,

$$\begin{aligned}
&\left(\frac{\tilde{N}}{B_{1p}(v) + B_{1u}(v) + B_{2p}(v) + B_{2u}(v)} \right) \xrightarrow{P} \frac{1}{[a_1\pi_{11} + \frac{a_1}{2}\pi_{10} + \frac{a_1}{2}\pi_{01} + a_2\pi_1 + a_3\pi_2]}, \\
&\left(\frac{B_{1p}(v) + B_{1u}(v)}{n+s} \right) \xrightarrow{P} \left[\left(\frac{a_1}{a_1+2a_2} \right) (\pi_{11} + \pi_{10}) + \left(\frac{2a_2}{a_1+2a_2} \right) \pi_1 \right], \text{ and} \\
&\left(\frac{B_{2p}(v) + B_{2u}(v)}{n+t} \right) \xrightarrow{P} \left[\left(\frac{a_1}{a_1+2a_3} \right) (\pi_{11} + \pi_{01}) + \left(\frac{2a_3}{a_1+2a_3} \right) \pi_2 \right].
\end{aligned}$$

Similarly,

$$\begin{aligned}
Q_{22} &= \sum_{j=1}^{\tilde{N}} \int_0^{t^*} H_{j2}^2(v) I(Z_j > v) d\Lambda(v) \\
&= \int_0^{t^*} \left(\frac{n+s}{\tilde{N}} \right) W^2(v) \left(\frac{\tilde{N}}{B_{1p}(v) + B_{1u}(v) + B_{2p}(v) + B_{2u}(v)} \right)^2 \times \\
&\quad \left(\frac{B_{1p}(v) + B_{1u}(v)}{n+s} \right)^2 \left(\frac{B_{2p}(v) + B_{2u}(v)}{n+t} \right) d\Lambda(v) \\
&\xrightarrow{P} \left(\frac{a_1 + 2a_2}{2} \right) \int_0^{t^*} W_*^2(v) \times \\
&\quad \frac{\left[\left(\frac{a_1}{a_1+2a_2} \right) (\pi_{11} + \pi_{10}) + \left(\frac{2a_2}{a_1+2a_2} \right) \pi_1 \right]^2 \left[\left(\frac{a_1}{a_1+2a_3} \right) (\pi_{11} + \pi_{01}) + \left(\frac{2a_3}{a_1+2a_3} \right) \pi_2 \right]}{[a_1\pi_{11} + \frac{a_1}{2}\pi_{10} + \frac{a_1}{2}\pi_{01} + a_2\pi_1 + a_3\pi_2]^2} d\Lambda(v).
\end{aligned}$$

$Q_{ii,\epsilon}$ can be shown to converge in probability to 0 using the proof of Theorem 7.2.1 in Fleming and Harrington.

So, Result A.1 obtains.

To this point, convergence over finite intervals, $[0, t^*]$, has been considered, but, as in Fleming in Harrington, we can extend the interval to $[0, \infty]$ for the statistics we consider, since the statistics satisfy the conditions of Corollary 7.2.1 of Fleming and Harrington. Thus, we obtain the following result.

Result A.2 Under H_0 ,

$$V' = U_1 - U_2 \xrightarrow{D} N(0, \sigma^2),$$

as $\tilde{N} \rightarrow \infty$, where

$$\begin{aligned} \sigma^2 &= \sigma_{11} + \sigma_{22} \\ &= \int_0^\infty W_*^2(v) \tilde{\pi}(v) d\Lambda(v), \end{aligned}$$

and

$$\tilde{\pi} = \frac{\left[\left(\frac{a_1}{a_1+2a_2}\right)(\pi_{11} + \pi_{10}) + \left(\frac{2a_2}{a_1+2a_2}\right)\pi_1\right]\left[\left(\frac{a_1}{a_1+2a_3}\right)(\pi_{11} + \pi_{01}) + \left(\frac{2a_3}{a_1+2a_3}\right)\pi_2\right]}{a_1\pi_{11} + \frac{a_1}{2}\pi_{10} + \frac{a_1}{2}\pi_{01} + a_2\pi_1 + a_3\pi_2}.$$

From Result A.2, we can obtain the asymptotic null distributions of V'_{PW} and V'_{AK} , for which the following lemma is needed.

Lemma A.1 Under H_0 , the weight functions $W(v)$ corresponding to V'_{PW} and V'_{AK} both converge in probability to $S(v)$, i.e.

$$\tilde{S}(v) \xrightarrow{P} S(v), \text{ and } \overline{\tilde{S}}(v) \xrightarrow{P} S(v),$$

as $\tilde{N} \rightarrow \infty$.

Proof. Convergence in probability of the Kaplan-Meier estimator in the presence of paired observations follows from Theorem 1 (see page 20) of Ying and Wei (1994).

Result A.3 Under H_0 , both V'_{PW} and V'_{AK} converge in distribution to a normal random variable with mean 0 and variance,

$$\sigma^2 = \int_0^\infty S^2(v) \tilde{\pi}(v) d\Lambda(v)$$

as $\tilde{N} \rightarrow \infty$.

Proof. Follows from Result A.2 and Lemma A.1.

An intuitively reasonable estimator of σ^2 is given by

$$\hat{\sigma}^2 = \frac{\tilde{N}}{(n+s)(n+t)} \int_0^\infty W^2(v) \frac{B_1(v)B_2(v)}{B^2(v)} dN(v).$$

We conjecture that unbiasedness and consistency of $\hat{\sigma}^2$ can be established along the lines of the proofs of Theorem 3.3.2 and Corollary 7.2.1 of Fleming and Harrington.

Note that in the two independent sample case ($n = 0$), the results given here correspond to those given by Fleming and Harrington.

APPENDIX B FORTRAN CODE

Following are listings of the Fortran programs used to calculate p-values for the *D-PW* and *D-AK* tests. To ensure correctness each program was developed by appropriately modifying and incorporating Edgington's (1987) tested programs. The programs follow the general algorithm:

1. Input observed data
2. Set gtcounter to 0 and set permcounter to 0
3. Compute test statistic
4. Add 1 to permcounter
5. If test statistic value is greater than or equal to observed test statistic value then add 1 to gtcounter
6. If permcounter equals the desired number of permutations to be used then go to 7; else go to 8
7. Calculate the p-value by dividing the gtcounter by the desired number of permutations to be used
8. Permute the data and go to 3

As listed, the programs compute one-sided p-values, but two-sided p-values can be obtained by slightly modifying the code in accordance with defining a two-sided rejection region. The dimensions given in the code can be adjusted as well to accomodate a given sample size.

We give five programs; one (Program 1) where the p-values are based on all $M = 2^n \times \binom{s+t}{s}$ permutations of the scores, and four (Programs 2 - 5) where the p-values are based on a random sample of $nperms = 100,000$ permutations ($nperms$ can be adjusted as needed). Since our testing procedures involve separately permuting the paired and unpaired observations, Programs 2 - 5 correspond to $2^n \leq [>] nperms$, and $\binom{s+t}{s} \leq [>] nperms$.

PROGRAM 1

```

PROGRAM MAIN
PARAMETER (NMAX=5000)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
DIMENSION X(NMAX),Y(NMAX),PSCOREX(NMAX),PSCOREY(NMAX),
1 ASCOREX(NMAX),ASCOREY(NMAX),
1 ATESTONE(100000),ATESTONEB(100000),
1 PTESTONE(100000),PTESTONEB(100000)
INTEGER DEATHX(NMAX),DEATHY(NMAX),S,T,NPAIRS,NN,
1 NNFAC,SFAC,TFAC,FACTOR1,FACTOR2,M
DOUBLE PRECISION MTEMP

C   INTRODUCE PROGRAM AND GET DATA FROM USER

      CALL INTRO(NPAIRS,S,T,NN,X,DEATHX,Y,DEATHY)

C   COMPUTE NUMBER OF PERMUTATIONS

      NNFAC=1
      SFAC=1
      TFAC=1

      DO 401 I=1,NN
401  NNFAC=NNFAC*I
      DO 402 I=1,S
402  SFAC=SFAC*I
      DO 403 I=1,T
403  TFAC=TFAC*I
      FAC1=NNFAC/(SFAC*TFAC)
      FACTOR1=INT(FAC1)

      FAC2=2**NPAIRS
      FACTOR2=INT(FAC2)

      MTEMP=FACTOR1*FACTOR2
      M=INT(MTEMP)

C   COMPUTE PRENTICE-WILCOXON TYPE SCORES

      CALL PW(X,DEATHX,Y,DEATHY,NPAIRS,S,T,PSCOREX,PSCOREY)

C   COMPUTE AKRITAS SCORES

      CALL AKRIT(X,DEATHX,Y,DEATHY,NPAIRS,S,T,ASCOREX,ASCOREY)

C   PERMUTE DATA AND COMPUTE P-VALUE

      CALL COMBO1(NPAIRS,S,T,NN,PSCOREX,PSCOREY,POBTONE,POBTONEB,
+ PTESTONE,PTESTONEB)
      CALL COMBO1(NPAIRS,S,T,NN,ASCOREX,ASCOREY,AOBTONE,AOBTONEB,
+ ATESTONE,ATESTONEB)

      CALL COMPUTE1(FACTOR1,FACTOR2,M,POBTONE,POBTONEB,PTESTONE,
+ PTESTONEB,PPVALUE,NPAIRS,S,T)

```

```
CALL COMPUTE1(FACTOR1,FACTOR2,M,AOBTONE,AOBTONEB,ATESTONE,
+ ATESTONEB,APVALUE,NPAIRS,S,T)
```

```
WRITE(6,*) ' '
WRITE(6,*) ' '
WRITE(6,1050) POBTONE
WRITE(6,1051) PPVALUE
WRITE(6,1052) M
WRITE(6,*) ' '
WRITE(6,*) ' '
WRITE(6,1053) AOBTONE
WRITE(6,1051) APVALUE
WRITE(6,1052) M
```

```
STOP
```

```
1050 FORMAT(' OBSERVED V-PW STATISTIC',F30.5)
```

```
1051 FORMAT(' P-VALUE',F8.6)
```

```
1052 FORMAT(' NO. PERMUTATIONS', I20)
```

```
1053 FORMAT(' OBSERVED V-AK STATISTIC',F30.5)
```

```
END
```

PROGRAM 2

```

PROGRAM MAIN
PARAMETER (NMAX=5000)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
DIMENSION X(NMAX),Y(NMAX),PSCOREX(NMAX),PSCOREY(NMAX),
1 ASCOREX(NMAX),ASCOREY(NMAX),
1 ATESTONE(100000),ATESTONEB(100000),
1 PTESTONE(100000),PTESTONEB(100000)
INTEGER DEATHX(NMAX),DEATHY(NMAX),S,T,NPAIRS,NN,
1 NNFAC,SFAC,TFAC,FACTOR1,FACTOR2,NPERMS

C   INTRODUCE PROGRAM AND GET DATA FROM USER

      CALL INTRO(NPAIRS,S,T,NN,X,DEATHX,Y,DEATHY)

C   COMPUTE NUMBER OF PERMUTATIONS

      NNFAC=1
      SFAC=1
      TFAC=1

      DO 401 I=1,NN
401  NNFAC=NNFAC*I
      DO 402 I=1,S
402  SFAC=SFAC*I
      DO 403 I=1,T
403  TFAC=TFAC*I
      FAC1=NNFAC/(SFAC*TFAC)
      FACTOR1=INT(FAC1)

      FAC2=2**NPAIRS
      FACTOR2=INT(FAC2)

      NPERMS=100000

C   COMPUTE PRENTICE-WILCOXON TYPE SCORES

      CALL PW(X,DEATHX,Y,DEATHY,NPAIRS,S,T,PSCOREX,PSCOREY)

C   COMPUTE AKRITAS SCORES

      CALL AKRIT(X,DEATHX,Y,DEATHY,NPAIRS,S,T,ASCOREX,ASCOREY)

C   PERMUTE DATA AND COMPUTE P-VALUE

      CALL COMBO1(NPAIRS,S,T,NN,PSCOREX,PSCOREY,POBTONE,POBTONEB,
+ PTESTONE,PTESTONEB)
      CALL COMBO1(NPAIRS,S,T,NN,ASCOREX,ASCOREY,AOBTONE,AOBTONEB,
+ ATESTONE,ATESTONEB)

      CALL COMPUTE2(FACTOR1,FACTOR2,NPERMS,POBTONE,POBTONEB,PTESTONE,
+ PTESTONEB,PPVALUE,NPAIRS,S,T)

      CALL COMPUTE2(FACTOR1,FACTOR2,NPERMS,AOBTONE,AOBTONEB,ATESTONE,
+ ATESTONEB,APVALUE,NPAIRS,S,T)

```

```
WRITE(6,*) ' '
WRITE(6,*) ' '
WRITE(6,1050) POBTONE
WRITE(6,1051) PPVALUE
WRITE(6,1052) NPERMS
WRITE(6,*) ' '
WRITE(6,*) ' '
WRITE(6,1053) AOBTONE
WRITE(6,1051) APVALUE
WRITE(6,1052) NPERMS

STOP
1050 FORMAT(' OBSERVED V-PW STATISTIC',F30.5)
1051 FORMAT(' P-VALUE',F8.6)
1052 FORMAT(' NO. PERMUTATIONS',I20)
1053 FORMAT(' OBSERVED V-AK STATISTIC',F30.5)
END
```

PROGRAM 3

```

PROGRAM MAIN
PARAMETER (NMAX=5000)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
DIMENSION X(NMAX),Y(NMAX),PSCOREX(NMAX),PSCOREY(NMAX),
1 ASCOREX(NMAX),ASCOREY(NMAX),
1 ATESTONE(100000),ATESTONEB(100000),
1 PTESTONE(100000),PTESTONEB(100000)
INTEGER DEATHX(NMAX),DEATHY(NMAX),S,T,NPAIRS,NN,
1 FACTOR2,NPERMS

C  INTRODUCE PROGRAM AND GET DATA FROM USER

      CALL INTRO(NPAIRS,S,T,NN,X,DEATHX,Y,DEATHY)

C  COMPUTE NUMBER OF PERMUTATIONS

      FAC2=2**NPAIRS
      FACTOR2=INT(FAC2)

      NPERMS=100000

C  COMPUTE PRENTICE-WILCOXON TYPE SCORES

      CALL PW(X,DEATHX,Y,DEATHY,NPAIRS,S,T,PSCOREX,PSCOREY)

C  COMPUTE AKRITAS SCORES

      CALL AKRIT(X,DEATHX,Y,DEATHY,NPAIRS,S,T,ASCOREX,ASCOREY)

C  PERMUTE DATA AND COMPUTE P-VALUE

      CALL COMBO2(NPAIRS,S,T,NN,PSCOREX,PSCOREY,POBTONE,POBTONEB,
+ PTESTONE,PTESTONEB,NPERMS)
      CALL COMBO2(NPAIRS,S,T,NN,ASCOREX,ASCOREY,AOBTONE,AOBTONEB,
+ ATESTONE,ATESTONEB,NPERMS)

      CALL COMPUTE3(NPERMS,FACTOR2,POBTONE,POBTONEB,PTESTONE,
+ PTESTONEB,PPVALUE,NPAIRS,S,T)

      CALL COMPUTE3(NPERMS,FACTOR2,AOBTONE,AOBTONEB,ATESTONE,
+ ATESTONEB,APVALUE,NPAIRS,S,T)

      WRITE(6,*) ' '
      WRITE(6,*) ' '
      WRITE(6,1050) POBTONE
      WRITE(6,1051) PPVALUE
      WRITE(6,1052) NPERMS
      WRITE(6,*) ' '
      WRITE(6,*) ' '
      WRITE(6,1053) AOBTONE
      WRITE(6,1051) APVALUE
      WRITE(6,1052) NPERMS

```

```
      STOP
1050 FORMAT(' OBSERVED V-PW STATISTIC',F30.5)
1051 FORMAT(' P-VALUE                                ',F8.6)
1052 FORMAT(' NO. PERMUTATIONS                        ', I20)
1053 FORMAT(' OBSERVED V-AK STATISTIC',F30.5)
      END
```

PROGRAM 4

```

PROGRAM MAIN
PARAMETER (NMAX=5000)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
DIMENSION X(NMAX),Y(NMAX),PSCOREX(NMAX),PSCOREY(NMAX),
1 ASOREX(NMAX),ASOREY(NMAX),
1 ATESTONE(100000),ATESTONEB(100000),
1 PTESTONE(100000),PTESTONEB(100000)
INTEGER DEATHX(NMAX),DEATHY(NMAX),S,T,NPAIRS,NN,
1 NPERMS,NNFAC,SFAC,TFAC,FACTOR1

C  INTRODUCE PROGRAM AND GET DATA FROM USER

      CALL INTRO(NPAIRS,S,T,NN,X,DEATHX,Y,DEATHY)

C  COMPUTE NUMBER OF PERMUTATIONS

      NNFAC=1
      SFAC=1
      TFAC=1

      DO 401 I=1,NN
401  NNFAC=NNFAC*I
      DO 402 I=1,S
402  SFAC=SFAC*I
      DO 403 I=1,T
403  TFAC=TFAC*I
      FAC1=NNFAC/(SFAC*TFAC)
      FACTOR1=INT(FAC1)

      NPERMS=100000

C  COMPUTE PRENTICE-WILCOXON TYPE SCORES

      CALL PW(X,DEATHX,Y,DEATHY,NPAIRS,S,T,PSCOREX,PSCOREY)

C  COMPUTE AKRITAS SCORES

      CALL AKRIT(X,DEATHX,Y,DEATHY,NPAIRS,S,T,ASOREX,ASOREY)

C  PERMUTE DATA AND COMPUTE P-VALUE

      CALL COMBO3(NPAIRS,S,T,NN,PSCOREX,PSCOREY,POBTONE,POBTONEB,
+ PTESTONE,PTESTONEB,NPERMS)
      CALL COMBO3(NPAIRS,S,T,NN,ASOREX,ASOREY,AOBTONE,AOBTONEB,
+ ATESTONE,ATESTONEB,NPERMS)

      CALL COMPUTE4(FACTOR1,NPERMS,POBTONE,POBTONEB,PTESTONE,
+ PTESTONEB,PPVALUE,NPAIRS,S,T)

      CALL COMPUTE4(FACTOR1,NPERMS,AOBTONE,AOBTONEB,ATESTONE,
+ ATESTONEB,APVALUE,NPAIRS,S,T)

      WRITE(6,*) ' '
      WRITE(6,*) ' '

```

```
WRITE(6,1050) POBTONE
WRITE(6,1051) PPVALUE
WRITE(6,1052) NPERMS
WRITE(6,*) ' '
WRITE(6,*) ' '
WRITE(6,1053) AOBTONE
WRITE(6,1051) APVALUE
WRITE(6,1052) NPERMS
```

```
STOP
```

```
1050 FORMAT(' OBSERVED V-PW STATISTIC',F30.5)
1051 FORMAT(' P-VALUE
1052 FORMAT(' NO. PERMUTATIONS      ', I20)
1053 FORMAT(' OBSERVED V-AK STATISTIC',F30.5)
END
```

```
',F8.6)
```

PROGRAM 5

```

PROGRAM MAIN
PARAMETER (NMAX=5000)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
DIMENSION X(NMAX),Y(NMAX),PSCOREX(NMAX),PSCOREY(NMAX),
1 ASCOREX(NMAX),ASCOREY(NMAX),
1 ATESTONE(100000),ATESTONEB(100000),
1 PTESTONE(100000),PTESTONEB(100000)
INTEGER DEATHX(NMAX),DEATHY(NMAX),S,T,NPAIRS,NN,
1 NPERMS

C  INTRODUCE PROGRAM AND GET DATA FROM USER

      CALL INTRO(NPAIRS,S,T,NN,X,DEATHX,Y,DEATHY)

C  COMPUTE NUMBER OF PERMUTATIONS

      NPERMS=100000

C  COMPUTE PRENTICE-WILCOXON TYPE SCORES

      CALL PW(X,DEATHX,Y,DEATHY,NPAIRS,S,T,PSCOREX,PSCOREY)

C  COMPUTE AKRITAS SCORES

      CALL AKRIT(X,DEATHX,Y,DEATHY,NPAIRS,S,T,ASCOREX,ASCOREY)

C  PERMUTE DATA AND COMPUTE P-VALUE

      CALL COMBO4(NPAIRS,S,T,NN,PSCOREX,PSCOREY,POBTONE,POBTONEB,
+ PTESTONE,PTESTONEB,NPERMS)
      CALL COMBO4(NPAIRS,S,T,NN,ASCOREX,ASCOREY,AOBTONE,AOBTONEB,
+ ATESTONE,ATESTONEB,NPERMS)

      CALL COMPUTE5(NPERMS,POBTONE,POBTONEB,PTESTONE,
+ PTESTONEB,PPVALUE,NPAIRS,S,T)

      CALL COMPUTE5(NPERMS,AOBTONE,AOBTONEB,ATESTONE,
+ ATESTONEB,APVALUE,NPAIRS,S,T)

      WRITE(6,*) ' '
      WRITE(6,*) ' '
      WRITE(6,1050) POBTONE
      WRITE(6,1051) PPVALUE
      WRITE(6,1052) NPERMS
      WRITE(6,*) ' '
      WRITE(6,*) ' '
      WRITE(6,1053) AOBTONE
      WRITE(6,1051) APVALUE
      WRITE(6,1052) NPERMS

      STOP
1050 FORMAT(' OBSERVED V-PW STATISTIC',F30.5)
1051 FORMAT(' P-VALUE',F8.6)

```

```
1052 FORMAT(' NO. PERMUTATIONS          ', I20)
1053 FORMAT(' OBSERVED V-AK STATISTIC',F30.5)
      END
```

SUBROUTINES

```

*****
*      SUBROUTINE INTRO
*
*      THIS SUBROUTINE INTRODUCES THE PROGRAM AND GETS DATA FROM USER
*
*****

SUBROUTINE INTRO(NPAIRS,S,T,NN,X,DEATHX,Y,DEATHY)

DOUBLE PRECISION AX(5000),BX(5000),AY(5000),BY(5000),
+ X(5000),Y(5000)
CHARACTER ANSWER*1,IFILE*60
INTEGER NPAIRS,S,T,NN,DEATHX(5000),DEATHY(5000),ENTRY,
+ I

WRITE(6,1001)
1001 FORMAT(/
1' THIS PROGRAM PERFORMS THE D-PW AND D-AK TESTS FOR SURVIVAL DATA'/
2' CONSISTING OF BOTH PAIRED AND UNPAIRED OBSERVATIONS.'/
3' THE PROGRAM ALSO WORKS FOR PAIRED OBSERVATIONS, AND FOR'/
4' TWO-SAMPLE (UNPAIRED) OBSERVATIONS.'//)
WRITE(6,*)' PRESS RETURN TO CONTINUE'
READ(5,1005) ANSWER
WRITE(6,1002)
1002 FORMAT(/
1' THE PROGRAM OUTPUTS THE VALUES OF THE TEST STATISTICS,'/
2' THE P-VALUES (ONE-SIDED), AND THE NUMBER OF PERMUTATIONS'/
3' USED TO COMPUTE THE P-VALUES.'//)
WRITE(6,*)' PRESS RETURN TO CONTINUE'
READ(5,1005) ANSWER
WRITE(6,1003)
1003 FORMAT(/
1' THE SURVIVAL DATA CAN BE ENTERED DIRECTLY ON THE'/
2' SCREEN, OR THROUGH A FILE. ON SCREEN ENTRY IS'/
3' MENU DRIVEN; IF ENTERING THROUGH A FILE, THE FORMAT IS:'/
4' X, INDICATOR FOR X, Y, INDICATOR FOR Y [THE PAIRED DATA]'/
5' X, INDICATOR FOR X [UNPAIRED X OBSERVATIONS]'/
6' Y, INDICATOR FOR Y [UNPAIRED Y OBSERVATIONS]'/
7' '/
8' X AND Y ARE SURVIVAL TIMES FOR WHICH UP TO 6 DECIMAL PLACES'/
9' MAY BE ENTERED. INDICATORS ARE 1 IF UNCENSORED, 0 IF CENSORED'/
9' THE X SAMPLE MUST BE THE SAMPLE HYPOTHESIZED TO'/
Z' HAVE THE LONGER SURVIVAL TIMES.'//)
WRITE(6,*)' PRESS RETURN TO CONTINUE'
READ(5,1005) ANSWER
WRITE(6,*)' ENTER 1 FOR ON-SCREEN ENTRY,'
WRITE(6,*)' ENTER 2 IF YOUR DATA IS ON FILE'
READ(5,*) ENTRY
IF (ENTRY.EQ.1) THEN
WRITE(6,*)' ENTER NUMBER OF PAIRS'
READ(5,*) NPAIRS
DO 200 I=1,NPAIRS
WRITE(6,*) ' ENTER X, INDICATOR FOR X, Y, INDICATOR FOR Y'
200 READ(5,*) X(I),DEATHX(I),Y(I),DEATHY(I)

```

```

WRITE(6,*) ' ENTER NUMBER OF UNPAIRED X OBSERVATIONS'
READ(5,*) S
DO 205 I=1,S
WRITE(6,*) ' ENTER X, INDICATOR FOR X'
205 READ(5,*) X(NPAIRS+I),DEATHX(NPAIRS+I)
WRITE(6,*) ' ENTER NUMBER OF UNPAIRED Y OBSERVATIONS'
READ(5,*) T
DO 210 I=1,T
WRITE(6,*) ' ENTER Y, INDICATOR FOR Y'
210 READ(5,*) Y(NPAIRS+I),DEATHY(NPAIRS+I)
ELSE
WRITE(6,*) ' ENTER NUMBER OF PAIRS'
READ(5,*) NPAIRS
WRITE(6,*) ' ENTER NUMBER OF UNPAIRED X OBSERVATIONS'
READ(5,*) S
WRITE(6,*) ' ENTER NUMBER OF UNPAIRED Y OBSERVATIONS'
READ(5,*) T
WRITE(6,*) ' ENTER NAME OF INPUT FILE'
READ(5,1006) IFILE
OPEN(8,FILE=IFILE)
DO 300 I=1,NPAIRS
300 READ(8,*) X(I),DEATHX(I),Y(I),DEATHY(I)
DO 305 I=1,S
305 READ(8,*) X(NPAIRS+I),DEATHX(NPAIRS+I)
DO 310 I=1,T
310 READ(8,*) Y(NPAIRS+I),DEATHY(NPAIRS+I)
CLOSE(8)
ENDIF

NN=S+T

C   KEEP 6 DECIMAL PLACES AND ADJUST CENSORED TIMES FOR SCORING

DO 320 I=1,(NPAIRS+S)
AX(I)=X(I)*1000000
BX(I)=INT(AX(I))
X(I)=BX(I)/1000000
320 IF(DEATHX(I).EQ.0) X(I)=X(I)+0.0000001
DO 330 I=1,(NPAIRS+T)
AY(I)=Y(I)*1000000
BY(I)=INT(AY(I))
Y(I)=BY(I)/1000000
330 IF(DEATHY(I).EQ.0) Y(I)=Y(I)+0.0000001
RETURN
1005 FORMAT(A1)
1006 FORMAT(A60)
END

*
*****
*
*   SUBROUTINE PW
*
*   THIS SUBROUTINE COMPUTES PRENTICE-WILCOXON TYPE SCORES
*
*****

SUBROUTINE PW(OBSX,DELTAX,OBSY,DELTAY,NPAIRS,S,T,PCSCOREX,PCSCOREY)

```

```

DOUBLE PRECISION PSCOREX(5000),PSCOREY(5000)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
DIMENSION OBSX(10000),OBSY(10000),SURV(0:20000)
DIMENSION OBS(20000),OBSS(20000)
INTEGER DELTAX(5000),DELTAY(5000),DELTA(20000),DTHS(20000),S,T
DO 9 I=1,(NPAIRS+S)
  OBS(I)=OBSX(I)
9  DELTA(I)=DELTAX(I)
  DO 10 I=1,(NPAIRS+T)
    OBS(NPAIRS+S+I)=OBSY(I)
10  DELTA(NPAIRS+S+I)=DELTAY(I)
  DO 12 I=1,(2*NPAIRS+S+T)
12  OBSS(I)=OBS(I)
    CALL SORTD(OBSS,1,(2*NPAIRS+S+T))
    DO 20 I=1,(2*NPAIRS+S+T)
      DO 15 J=1,(2*NPAIRS+S+T)
        IF(OBSS(I).EQ. OBS(J)) THEN
          DTHS(I)=DELTA(J)
          GO TO 16
        ENDIF
15  CONTINUE
16  CONTINUE
20  CONTINUE
    CALL KM2(OBSS,DTHS,SURV,(2*NPAIRS+S+T))
    DO 30 I=1,(NPAIRS+S)
      CALL PLACE(OBSX(I),OBSS,(2*NPAIRS+S+T),IPLACX)
      F=SURV(IPLACX)
      IF(DELTAX(I).EQ. 1) THEN
        PSCOREX(I)=1.0 - 2.0*F
      ELSE
        PSCOREX(I)=1.0 - F
      ENDIF
30  CONTINUE
    DO 40 I=1,(NPAIRS+T)
      CALL PLACE(OBSY(I),OBSS,(2*NPAIRS+S+T),IPLACY)
      F=SURV(IPLACY)
      IF(DELTAY(I).EQ. 1) THEN
        PSCOREY(I)=1.0 - 2.0*F
      ELSE
        PSCOREY(I)=1.0 - F
      ENDIF
40  CONTINUE
    RETURN
  END

```

```

*
*****
*
*      SUBROUTINE AKRIT
*
*      THIS SUBROUTINE COMPUTES AKRITAS SCORES
*
*      BASED ON CODE SUPPLIED BY DR. THOMAS O'GORMAN
*
*****

```

```

SUBROUTINE AKRIT(OBSX,DEATHX,OBSY,DEATHY,NPAIRS,S,T,
+ ASCOREX,ASCOREY)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)

```

```

DIMENSION OBSX(5000),OBSY(5000),SX(0:5000),SY(0:5000)
DIMENSION ASCOREX(5000),ASCOREY(5000),OBSXS(10000),OBSYS(10000)
INTEGER DEATHX(5000),DEATHY(5000),DTHXS(10000),DTHYS(10000),S,T,
1 NPAIRS
DO 9 I=1,(NPAIRS+S)
9 OBSXS(I)=OBSX(I)
DO 10 I=1,(NPAIRS+T)
10 OBSYS(I)=OBSY(I)
CALL SORTD(OBSXS,1,(NPAIRS+S))
CALL SORTD(OBSYS,1,(NPAIRS+T))
DO 20 I=1,(NPAIRS+S)
DO 15 J=1,(NPAIRS+S)
IF(OBSXS(I) .EQ. OBSX(J)) THEN
    DTHXS(I)=DEATHX(J)
    GOTO 20
ENDIF
15 CONTINUE
20 CONTINUE
DO 22 I=1,(NPAIRS+T)
DO 21 J=1,(NPAIRS+T)
IF(OBSYS(I) .EQ. OBSY(J)) THEN
    DTHYS(I)=DEATHY(J)
    GOTO 22
ENDIF
21 CONTINUE
22 CONTINUE
CALL KM(OBSXS,DTHXS,SX,(NPAIRS+S))
CALL KM(OBSYS,DTHYS,SY,(NPAIRS+T))
DO 30 I=1,(NPAIRS+S)
CALL PLACE(OBSX(I),OBSXS,(NPAIRS+S),IPLACX)
CALL PLACE(OBSY(I),OBSYS,(NPAIRS+T),IPLACY)
F=((1.0-SX(IPLACX))+(1.0-SY(IPLACY)))/2.0
IF(DEATHX(I) .EQ. 1) THEN
    ASCOREX(I)=F
ELSE
    ASCOREX(I)=(1.0+F)/2.0
ENDIF
30 CONTINUE
DO 40 I=1,(NPAIRS+T)
CALL PLACE(OBSY(I),OBSXS,(NPAIRS+S),IPLACX)
CALL PLACE(OBSY(I),OBSYS,(NPAIRS+T),IPLACY)
F=((1.0-SX(IPLACX))+(1.0-SY(IPLACY)))/2.0
IF(DEATHY(I) .EQ. 1) THEN
    ASCOREY(I)=F
ELSE
    ASCOREY(I)=(1.0+F)/2.0
ENDIF
40 CONTINUE
RETURN
END

```

```

*
*****
*
* SUBROUTINE SORTD
*
* THIS SUBROUTINE SORTS AN ARRAY OF DOUBLE PRECISION FLOATING
* POINT NUMBERS.
*

```

CODE SUPPLIED BY DR. THOMAS O'GORMAN

SUBROUTINE SORTD(A,II,JJ)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)

```

C
C THIS IS THE DOUBLE PRECISION VERSION OF SORT
C
C SORTS ARRAY A INTO INCREASING ORDER, FROM A(II) TO A(JJ)
C ORDERING IS BY SUBTRACTION
C ARRAYS IU(K) AND IL(K) PERMIT SORTING UP TO 2**(K+1)-1 ELEMENTS
  INTEGER IU(16),IL(16)
  DOUBLE PRECISION A(1)
  M=1
  I=II
  J=JJ
  5 IF(I .GE. J) GO TO 70
10 K=I
  IJ=(J+I)/2
  T=A(IJ)
  IF(A(I) .LE. T) GO TO 20
  A(IJ)=A(I)
  A(I)=T
  T=A(IJ)
20 L=J
  IF(A(J) .GE. T) GO TO 40
  A(IJ)=A(J)
  A(J)=T
  T=A(IJ)
  IF(A(I) .LE. T) GO TO 40
  A(IJ)=A(I)
  A(I)=T
  T=A(IJ)
  GO TO 40
30 A(L)=A(K)
  A(K)=TT
40 L=L-1
  IF(A(L) .GT. T) GO TO 40
  TT=A(L)
50 K=K+1
  IF(A(K) .LT. T) GO TO 50
  IF(K .LE. L) GO TO 30
  IF(L-I .LE. J-K) GO TO 60
  IL(M)=I
  IU(M)=L
  I=K
  M=M+1
  GO TO 80
60 IL(M)=K
  IU(M)=J
  J=L
  M=M+1
  GO TO 80
70 M=M-1
  IF(M .EQ. 0) RETURN
  I=IL(M)
  J=IU(M)
80 IF(J-I .GE. 11) GO TO 10
  IF(I .EQ. 11) GO TO 5

```

```

      I=I-1
90   I=I+1
      IF(I .EQ. J) GO TO 70
      T=A(I+1)
      IF(A(I) .LE. T) GO TO 90
      K=I
100  A(K+1)=A(K)
      K=K-1
      IF(T .LT. A(K)) GO TO 100
      A(K+1)=T
      GO TO 90
      END

```

```

*
*****
*
*      SUBROUTINE PLACE
*
*      THIS SUBROUTINE RETURNS THE INDEX WHERE A NUMBER
*      WOULD BE PLACED IN A SORTED ARRAY.
*
*      CODE SUPPLIED BY DR. THOMAS O'GORMAN
*
*****
*

```

```

      SUBROUTINE PLACE (X,T,N,IPLACE)
C
C      THIS SUBROUTINE RETURNS THE INDEX (IPLACE) WHERE
C      T(IPLACE) .LE. X .LE. T(IPLACE + 1) .
C
C      THE ARRAY T MUST BE SORTED.
C
      DOUBLE PRECISION X,T(N)
      IF (X .LT. T(1)) THEN
          IPLACE=0
          RETURN
      ENDIF
      DO 100 I=2,N
          IF(X .LT. T(I)) THEN
              IPLACE= I-1
              RETURN
          ENDIF
100  CONTINUE
      IPLACE = N
      RETURN
      END

```

```

*
*****
*
*      SUBROUTINE KM
*
*      THIS SUBROUTINE COMPUTES KAPLAN-MEIER SURVIVAL ESTIMATES.
*
*      CODE SUPPLIED BY DR. THOMAS O'GORMAN
*
*****
*

```

```

      TIME MUST BE SORTED
      DELTA IS CENSORING INDICATOR 1=UNCENSORED 0 =CENSORED
      SURV IS PRODUCT LIMIT SURVIVAL ESTIMATES (OUTPUT)
      NUM IS NUMBER OF POINTS
C
C      SUBROUTINE KM(TIME,DELTA,SURV,NUM)
C      IMPLICIT DOUBLE PRECISION (A-H,O-Z)

```

```

      DIMENSION TIME(NUM),SURV(0:NUM)
      INTEGER DELTA(NUM)
      DOUBLE PRECISION N
      N=NUM
      SURV(0)=1.0
      DO 100 I=1,NUM
        IM1=I-1
        IF (DELTA(I) .EQ. 1) THEN
          SURV(I)=SURV(IM1)*(N-DBLE(I))/(N-DBLE(I)+1.0)
        ELSE
          SURV(I)=SURV(IM1)
        ENDIF
100    CONTINUE
      RETURN
      END
*****
*
*      SUBROUTINE KM2
*
*      THIS SUBROUTINE COMPUTES KAPLAN-MEIER TYPE SURVIVAL ESTIMATES.
*
*      BASED ON KM SUBROUTINE
*****
*
C      TIME MUST BE SORTED
C      DELTA IS CENSORING INDICATOR 1=UNCENSORED 0 =CENSORED
C      SURV IS PRODUCT LIMIT SURVIVAL ESTIMATES (OUTPUT)
C      NUM IS NUMBER OF POINTS
C
      SUBROUTINE KM2(TIME,DELTA,SURV,NUM)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION TIME(NUM),SURV(0:NUM)
      INTEGER DELTA(NUM)
      DOUBLE PRECISION N
      N=NUM+DBLE(1)
      SURV(0)=1.0
      DO 100 I=1,NUM
        IM1=I-1
        IF (DELTA(I) .EQ. 1) THEN
          SURV(I)=SURV(IM1)*(N-DBLE(I))/(N-DBLE(I)+1.0)
        ELSE
          SURV(I)=SURV(IM1)
        ENDIF
100    CONTINUE
      RETURN
      END
*****
*
*      SUBROUTINE COMBO1
*
*      THIS SUBROUTINE PERMUTES THE DATA
*
*      BASED ON CODE FOUND IN EDGINGTON (1987)
*****
*
      SUBROUTINE COMBO1(NPAIRS,S,T,NN,SCOREX,SCOREY,OBTONE,OBTONEB,
+ TESTONE,TESTONEB)

```

```

    INTEGER NTRE,ITRE,I,J,INDEX(5000,2),WITHOUT(5000,2),NP,
+   INDEXB(10000),NPAIRS,S,T,NN,QU
    DOUBLE PRECISION TOTAL,DATA(5000,2),TEST,SUBTOT,TESTONE(100000),
+   SUM,OBTONE,TOTALB,DATAB(10000),SUMB,
+   OBTONEB,TESTONEB(100000),SCOREX(5000),SCOREY(5000)

```

```

    NTRE=2
    TOTAL=0
    NP=1
    ITRE=NTRE-1
    QU=2

```

```

    IF (NPAIRS.EQ.0) GO TO 95
    DO 7 I=1,NPAIRS
      DATA(I,1)=SCOREX(I)
7     DATA(I,2)=SCOREY(I)
    DO 10 I=1,NPAIRS
      DO 10 J=1,NTRE
        TOTAL=TOTAL+DATA(I,J)
        INDEX(I,J)=J
        WITHOUT(I,J)=0
10     IF (J.EQ.NTRE)WITHOUT(I,J)=1
20     TEST=0
        SUBTOT=0
        TESTONE(NP)=0
        DO 40 J=1,ITRE
          SUM=0
          DO 30 I=1,NPAIRS
30         SUM=SUM+DATA(I,INDEX(I,J))
          SUBTOT=SUBTOT+SUM
40         TEST=TEST+SUM
          TEST=TEST+(TOTAL-SUBTOT)
          TESTONE(NP)=TESTONE(NP)+SUM
          NP=NP+1
          I=NPAIRS
50         J=ITRE
60         WITHOUT(I,INDEX(I,J))=1
70         IF (INDEX(I,J).EQ.NTRE)GO TO 80
          INDEX(I,J)=INDEX(I,J)+1
          IF (WITHOUT(I,INDEX(I,J)).EQ.0)GO TO 70
          WITHOUT(I,INDEX(I,J))=0
          IF (J.EQ.ITRE)GO TO 20
          J=J+1
          INDEX(I,J)=0
          GO TO 70
80         J=J-1
          IF (J.GT.0)GO TO 60
          DO 90 J=1,NTRE
            INDEX(I,J)=J
90         WITHOUT(I,J)=0
            WITHOUT(I,NTRE)=1
            I=I-1
            IF (I.NE.0)GO TO 50
            OBTONE=TESTONE(1)

```

C UNPAIRED DATA

```

      IF ((S.EQ.0).AND.(T.EQ.0)) THEN
        GO TO 900
      ELSE IF ((S.GT.0).AND.(T.EQ.0)) THEN
        TOTALB=0
        DO 91 I=1,S
91      TOTALB=TOTALB+SCOREX(NPAIRS+I)
        OBTONEB=TOTALB
      ELSE IF ((S.EQ.0).AND.(T.GT.0)) THEN
        TOTALB=0
        DO 92 I=1,T
92      TOTALB=TOTALB+SCOREY(NPAIRS+I)
        OBTONEB=TOTALB
      ELSE
95      TOTALB=0
        DO 100 I=1,S
100     DATAB(I)=SCOREX(NPAIRS+I)
        DO 101 I=1,T
101     DATAB(S+I)=SCOREY(NPAIRS+I)
        DO 102 I=1,NN
          TOTALB=TOTALB+DATAB(I)
          IF(I.EQ.S)SUMB=TOTALB
102     INDEXB(I)=I
          OBTONEB=SUMB
          TESTONEB(1)=OBTONEB
200     I=S
300     IF(INDEXB(I).EQ.NN)GO TO 600
          INDEXB(I)=INDEXB(I)+1
          IF(I.EQ.S)GO TO 400
          I=I+1
          INDEXB(I)=INDEXB(I-1)
          GO TO 300
400     SUMB=0
        DO 500 I=1,S
500     SUMB=SUMB+DATAB(INDEXB(I))
          TESTONEB(QU)=SUMB
          QU=QU+1
          GO TO 200
600     I=I-1
          IF(I.NE.0)GO TO 300
        ENDIF
900    CONTINUE
        RETURN
      END
*****
*
*      SUBROUTINE COMBO2
*
*      THIS SUBROUTINE PERMUTES THE DATA
*
*      BASED ON CODE FOUND IN EDGINGTON (1987)
*
*****

      SUBROUTINE COMBO2(NPAIRS,S,T,NN,SCOREX,SCOREY,OBTONE,OBTONEB,
+ TESTONE,TESTONEB,NPERMS)

      INTEGER NTRE,ITRE,I,J,INDEX(50,2),WITHOUT(50,2),NP,
+ NPAIRS,S,T,NN,K,QU
      DOUBLE PRECISION TOTAL,DATA(50,2),TEST,SUBTOT,TESTONE(5000),
+ SUM,OBTONE,TOTALB,DATAB(50),SUMB,XUNP,

```

```
+ OBTONEB,TESTONEB(5000),SCOREX(100),
+ SCOREY(100)
```

```
NTRE=2
TOTAL=0
NP=1
ITRE=NTRE-1
QU=2
```

```
IF (NPAIRS.EQ.0) GO TO 95
DO 7 I=1,NPAIRS
DATA(I,1)=SCOREX(I)
7 DATA(I,2)=SCOREY(I)
DO 10 I=1,NPAIRS
DO 10 J=1,NTRE
TOTAL=TOTAL+DATA(I,J)
INDEX(I,J)=J
WITHOUT(I,J)=0
10 IF (J.EQ.NTRE)WITHOUT(I,J)=1
20 TEST=0
SUBTOT=0
TESTONE(NP)=0
DO 40 J=1,ITRE
SUM=0
DO 30 I=1,NPAIRS
30 SUM=SUM+DATA(I,INDEX(I,J))
SUBTOT=SUBTOT+SUM
40 TEST=TEST+SUM
TEST=TEST+(TOTAL-SUBTOT)
TESTONE(NP)=TESTONE(NP)+SUM
NP=NP+1
I=NPAIRS
50 J=ITRE
60 WITHOUT(I,INDEX(I,J))=1
70 IF (INDEX(I,J).EQ.NTRE)GO TO 80
INDEX(I,J)=INDEX(I,J)+1
IF (WITHOUT(I,INDEX(I,J)).EQ.0)GO TO 70
WITHOUT(I,INDEX(I,J))=0
IF (J.EQ.ITRE)GO TO 20
J=J+1
INDEX(I,J)=0
GO TO 70
80 J=J-1
IF (J.GT.0)GO TO 60
DO 90 J=1,NTRE
INDEX(I,J)=J
90 WITHOUT(I,J)=0
WITHOUT(I,NTRE)=1
I=I-1
IF (I.NE.0)GO TO 50
OBTONE=TESTONE(1)
```

C UNPAIRED DATA

```
IF ((S.EQ.0).AND.(T.EQ.0)) THEN
GO TO 900
ELSE IF ((S.GT.0).AND.(T.EQ.0)) THEN
TOTALB=0
```

```

DO 91 I=1,S
91  TOTALB=TOTALB+SCOREX(NPAIRS+I)
    OBTONEB=TOTALB
    ELSE IF ((S.EQ.0).AND.(T.GT.0)) THEN
        TOTALB=0
        DO 92 I=1,T
92  TOTALB=TOTALB+SCOREY(NPAIRS+I)
    OBTONEB=TOTALB
    ELSE
95  TOTALB=0
    DO 100 I=1,S
100  DATAB(I)=SCOREX(NPAIRS+I)
    DO 101 I=1,T
101  DATAB(S+I)=SCOREY(NPAIRS+I)
    DO 102 I=1,NN
        TOTALB=TOTALB+DATAB(I)
102  IF(I.EQ.S)SUMB=TOTALB
    OBTONEB=SUMB
    TESTONEB(1)=OBTONEB
    DO 300 I=2,NPERMS
        SUMB=0
        DO 200 J=1,S
            K=J+RANF(0)*(NN-J+1)
            XUNP=DATAB(K)
            DATAB(K)=DATAB(J)
            DATAB(J)=XUNP
200  SUMB=SUMB+XUNP
        TESTONEB(I)=SUMB
300  CONTINUE
    ENDIF
900  CONTINUE
    RETURN
    END
*****
*
*  SUBROUTINE COMB03
*
*  THIS SUBROUTINE PERMUTES THE DATA
*
*  BASED ON CODE FOUND IN EDGINGTON (1987)
*
*****

SUBROUTINE COMB03(NPAIRS,S,T,NN,SCOREX,SCOREY,OBTONE,OBTONEB,
+ TESTONE,TESTONEB,NPERMS)

INTEGER NTRE,ITRE,I,J,
+ NPAIRS,S,T,NN,K,NPERMS,ID,INDEXB(50),QU
DOUBLE PRECISION TOTAL,DATA(50,2),TEST,SUBTOT,
+ TESTONE(5000),SUM,
+ OBTONE,TOTALB,DATAB(50),SUMB,XPAI,
+ OBTONEB,TESTONEB(5000),
+ SCOREX(100),SCOREY(100)

NTRE=2
TOTAL=0
ITRE=NTRE-1
NPERMS=100000
QU=2

```

```

      IF (NPAIRS.EQ.0) GO TO 95
      DO 7 I=1,NPAIRS
      DATA(I,1)=SCOREX(I)
7     DATA(I,2)=SCOREY(I)
      DO 10 I=1,NPAIRS
      DO 10 J=1,NTRE
10    TOTAL=TOTAL+DATA(I,J)
      DO 40 I=1,NPERMS
      TEST=0
      SUBTOT=0
      TESTONE(1)=0
      DO 30 K=1,ITRE
      SUM=0
      DO 20 J=1,NPAIRS
      IF(I.EQ.1) GO TO 20
      ID=K+RANF(0)*(NTRE-K+1)
      XPAI=DATA(J,K)
      DATA(J,K)=DATA(J,ID)
      DATA(J,ID)=XPAI
20    SUM=SUM+DATA(J,K)
      SUBTOT=SUBTOT+SUM
30    TEST=TEST+SUM
      TEST=TEST+(TOTAL-SUBTOT)
      TESTONE(I)=TESTONE(I)+SUM
40    CONTINUE
      OBTONE=TESTONE(1)

```

C UNPAIRED DATA

```

      IF ((S.EQ.0).AND.(T.EQ.0)) THEN
      GO TO 900
      ELSE IF ((S.GT.0).AND.(T.EQ.0)) THEN
      TOTALB=0
      DO 91 I=1,S
91    TOTALB=TOTALB+SCOREX(NPAIRS+I)
      OBTONEB=TOTALB
      ELSE IF ((S.EQ.0).AND.(T.GT.0)) THEN
      TOTALB=0
      DO 92 I=1,T
92    TOTALB=TOTALB+SCOREY(NPAIRS+I)
      OBTONEB=TOTALB
      ELSE
95    TOTALB=0
      DO 100 I=1,S
100   DATAB(I)=SCOREX(NPAIRS+I)
      DO 101 I=1,T
101   DATAB(S+I)=SCOREY(NPAIRS+I)
      DO 102 I=1,NN
      TOTALB=TOTALB+DATAB(I)
      IF(I.EQ.S)SUMB=TOTALB
102   INDEXB(I)=I
      OBTONEB=SUMB
      TESTONEB(1)=OBTONEB
200   I=S
300   IF(INDEXB(I).EQ.NN)GO TO 600
      INDEXB(I)=INDEXB(I)+1
      IF(I.EQ.S)GO TO 400
      I=I+1

```

```

        INDEXB(I)=INDEXB(I-1)
        GO TO 300
400    SUMB=0
        DO 500 I=1,S
500    SUMB=SUMB+DATAB(INDEXB(I))
        TESTONEB(QU)=SUMB
        QU=QU+1
        GO TO 200
600    I=I-1
        IF(I.NE.0)GO TO 300
        ENDIF
900    CONTINUE
        RETURN
        END

```

```

*****
*
*   SUBROUTINE COMBO4
*
*   THIS SUBROUTINE PERMUTES THE DATA
*
*   BASED ON CODE FOUND IN EDGINGTON (1987)
*
*****

```

```

SUBROUTINE COMBO4(NPAIRS,S,T,NN,SCOREX,SCOREY,OBTONE,OBTONEB,
+ TESTONE,TESTONEB,NPERMS)

```

```

    INTEGER NTRE,ITRE,I,J,NPAIRS,S,T,NN,K,NPERMS,ID
    DOUBLE PRECISION TOTAL,DATA(50,2),TEST,SUBTOT,
+ TESTONE(5000),SUM,
+ OBTONE,TOTALB,DATAB(50),SUMB,XPAI,
+ OBTONEB,TESTONEB(5000),XUNP,
+ SCOREX(100),SCOREY(100)

```

```

    NTRE=2
    TOTAL=0
    ITRE=NTRE-1
    NPERMS=100000

```

```

    IF (NPAIRS.EQ.0) GO TO 95
    DO 7 I=1,NPAIRS
        DATA(I,1)=SCOREX(I)
7    DATA(I,2)=SCOREY(I)
        DO 10 I=1,NPAIRS
            DO 10 J=1,NTRE
10    TOTAL=TOTAL+DATA(I,J)
        DO 40 I=1,NPERMS
            TEST=0
            SUBTOT=0
            TESTONE(I)=0
            DO 30 K=1,ITRE
                SUM=0
                DO 20 J=1,NPAIRS
                    IF(I.EQ.1) GO TO 20
                    ID=K+RANF(0)*(NTRE-K+1)
                    XPAI=DATA(J,K)
                    DATA(J,K)=DATA(J,ID)
                    DATA(J,ID)=XPAI
20    SUM=SUM+DATA(J,K)

```

```

SUBTOT=SUBTOT+SUM
30  TEST=TEST+SUM
    TEST=TEST+(TOTAL-SUBTOT)
    TESTONE(I)=TESTONE(I)+SUM
40  CONTINUE
    OBTONE=TESTONE(1)

```

C UNPAIRED DATA

```

    IF ((S.EQ.0).AND.(T.EQ.0)) THEN
        GO TO 900
    ELSE IF ((S.GT.0).AND.(T.EQ.0)) THEN
        TOTALB=0
        DO 91 I=1,S
91     TOTALB=TOTALB+SCOREX(NPAIRS+I)
        OBTONEB=TOTALB
    ELSE IF ((S.EQ.0).AND.(T.GT.0)) THEN
        TOTALB=0
        DO 92 I=1,T
92     TOTALB=TOTALB+SCOREY(NPAIRS+I)
        OBTONEB=TOTALB
    ELSE
95     TOTALB=0
        DO 100 I=1,S
100    DATAB(I)=SCOREX(NPAIRS+I)
        DO 101 I=1,T
101    DATAB(S+I)=SCOREY(NPAIRS+I)
        DO 102 I=1,NN
            TOTALB=TOTALB+DATAB(I)
102    IF(I.EQ.S)SUMB=TOTALB
        OBTONEB=SUMB
        TESTONEB(1)=OBTONEB
        DO 300 I=2,NPERMS
            SUMB=0
            DO 200 J=1,S
                K=J+RANF(0)*(NN-J+1)
                XUNP=DATAB(K)
                DATAB(K)=DATAB(J)
                DATAB(J)=XUNP
200    SUMB=SUMB+XUNP
            TESTONEB(I)=SUMB
300    CONTINUE
        ENDIF
900    CONTINUE
        RETURN
        END

```

```

*
*****

```

```

*
* SUBROUTINE COMPUTE1
*
* THIS SUBROUTINE COMPUTES THE P-VALUE
*

```

```

*****

```

```

SUBROUTINE COMPUTE1(FACTOR1,FACTOR2,M,OBTONE,OBTONEB,TESTONE,
+ TESTONEB,PVALUE,NPAIRS,S,T)

```

```

INTEGER I,J,NGE,FACTOR1,FACTOR2,M,NPAIRS,S,T,K
DOUBLE PRECISION PERMSTAT(100000),OBTONE,OBTONEB,TESTONE(100000),

```

```
+ TESTONEB(100000),OBSTAT,PVALUE
```

```
NGE=0
```

```

IF (NPAIRS.EQ.0) THEN
  DO 10 I=1,FACTOR1
    IF (TESTONEB(I).GE.OBTONEB) NGE=NGE+1
10  CONTINUE
    GOTO 200
  ELSE IF ((S.EQ.0).AND.(T.EQ.0)) THEN
    DO 15 I=1,FACTOR2
      IF (TESTONE(I).GE.OBTONE) NGE=NGE+1
15  CONTINUE
      GOTO 200
    ELSE IF ((S.GT.0).AND.(T.EQ.0)) THEN
      OBSTAT=OBTONE+OBTONEB
      DO 16 I=1,FACTOR2
        PERMSTAT(I)=TESTONE(I)+OBTONEB
        IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
16  CONTINUE
        GOTO 200
      ELSE IF ((S.EQ.0).AND.(T.GT.0)) THEN
        OBSTAT=OBTONE
        DO 17 I=1,FACTOR2
          PERMSTAT(I)=TESTONE(I)
          IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
17  CONTINUE
          GOTO 200
        ELSE
          K=1
          OBSTAT=OBTONE+OBTONEB
          DO 20 I=1,FACTOR2
            DO 25 J=1,FACTOR1
              PERMSTAT(K)=TESTONE(I)+TESTONEB(J)
              K=K+1
25  CONTINUE
20  CONTINUE
          DO 50 I=1,M
            IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
50  CONTINUE
          END IF
200 PVALUE=DBLE(NGE)/M
    RETURN
  END

```

```

*
*****
*
*   SUBROUTINE COMPUTE2
*
*   THIS SUBROUTINE COMPUTES THE P-VALUE
*
*****

```

```

SUBROUTINE COMPUTE2(FACTOR1,FACTOR2,NPERMS,OBTONE,OBTONEB,TESTONE,
+ TESTONEB,PVALUE,NPAIRS,S,T)

```

```

  INTEGER I,NGE,FACTOR1,FACTOR2,NPERMS,NPAIRS,S,T,H,L
  DOUBLE PRECISION PERMSTAT(100000),OBTONE,OBTONEB,TESTONE(100000),
+ TESTONEB(100000),OBSTAT,PVALUE

```

```

NGE=0

IF (NPAIRS.EQ.0) THEN
  DO 10 I=1,FACTOR1
    IF (TESTONEB(I).GE.OBTONEB) NGE=NGE+1
10  CONTINUE
    GOTO 200
  ELSE IF ((S.EQ.0).AND.(T.EQ.0)) THEN
    DO 15 I=1,FACTOR2
      IF (TESTONE(I).GE.OBTONE) NGE=NGE+1
15  CONTINUE
      GOTO 200
    ELSE IF ((S.GT.0).AND.(T.EQ.0)) THEN
      OBSTAT=OBTONE+OBTONEB
      DO 16 I=1,FACTOR2
        PERMSTAT(I)=TESTONE(I)+OBTONEB
        IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
16  CONTINUE
        GOTO 200
      ELSE IF ((S.EQ.0).AND.(T.GT.0)) THEN
        OBSTAT=OBTONE
        DO 17 I=1,FACTOR2
          PERMSTAT(I)=TESTONE(I)
          IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
17  CONTINUE
          GOTO 200
        ELSE
          OBSTAT=OBTONE+OBTONEB
          PERMSTAT(1)=OBSTAT
          DO 37 I=2,NPERMS
            H=IGNUIN(1,FACTOR2)
            L=IGNUIN(1,FACTOR1)
            PERMSTAT(I)=TESTONE(H)+TESTONEB(L)
37  CONTINUE
            DO 50 I=1,NPERMS
              IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
50  CONTINUE
            END IF
          200 PVALUE=DBLE(NGE)/NPERMS
            RETURN
            END
*
*****
*
*   SUBROUTINE COMPUTE3
*
*   THIS SUBROUTINE COMPUTES THE P-VALUE
*
*****

SUBROUTINE COMPUTE3(NPERMS,FACTOR2,OBTONE,OBTONEB,TESTONE,
+ TESTONEB,PVALUE,NPAIRS,S,T)

INTEGER I,NGE,FACTOR2,NPERMS,NPAIRS,S,T,H
DOUBLE PRECISION PERMSTAT(100000),OBTONE,OBTONEB,TESTONE(100000),
+ TESTONEB(100000),OBSTAT,PVALUE

```

NGE=0

```

IF (NPAIRS.EQ.0) THEN
  PERMSTAT(1)=OBTONEB
  DO 10 I=2,NPERMS
    PERMSTAT(I)=TESTONEB(I)
10  CONTINUE
  DO 20 I=1,NPERMS
    IF (PERMSTAT(I).GE.OBTONEB) NGE=NGE+1
20  CONTINUE
    GOTO 200
  ELSE
    OBSTAT=OBTONE+OBTONEB
    PERMSTAT(1)=OBSTAT
    DO 37 I=2,NPERMS
      H=IGNUIN(1,FACTOR2)
      PERMSTAT(I)=TESTONE(H)+TESTONEB(I)
37  CONTINUE
      DO 50 I=1,NPERMS
        IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
50  CONTINUE
      END IF
200 PVALUE=DBLE(NGE)/NPERMS
    RETURN
    END

```

*

*
* SUBROUTINE COMPUTE4

*
* THIS SUBROUTINE COMPUTES THE P-VALUE

*

```

SUBROUTINE COMPUTE4(FACTOR1,NPERMS,OBTONE,OBTONEB,TESTONE,
+ TESTONEB,PVALUE,NPAIRS,S,T)

```

```

  INTEGER I,NGE,FACTOR1,NPERMS,NPAIRS,S,T,H
  DOUBLE PRECISION PERMSTAT(100000),OBTONE,OBTONEB,TESTONE(100000),
+ TESTONEB(100000),OBSTAT,PVALUE

```

NGE=0

```

IF ((S.EQ.0).AND.(T.EQ.0)) THEN
  DO 10 I=1,NPERMS
    IF (TESTONE(I).GE.OBTONE) NGE=NGE+1
10  CONTINUE
    GOTO 200
  ELSE IF ((S.GT.0).AND.(T.EQ.0)) THEN
    OBSTAT=OBTONE+OBTONEB
    DO 16 I=1,NPERMS
      PERMSTAT(I)=TESTONE(I)+OBTONEB
      IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
16  CONTINUE
      GOTO 200
  ELSE IF ((S.EQ.0).AND.(T.GT.0)) THEN
    OBSTAT=OBTONE
    DO 17 I=1,NPERMS

```

```

      PERMSTAT(I)=TESTONE(I)
      IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
17  CONTINUE
      GOTO 200
    ELSE
      OBSTAT=OBTONE+OBTONEB
      PERMSTAT(1)=OBSTAT
      DO 37 I=2,NPERMS
        H=IGNUIN(I,FACTOR1)
        PERMSTAT(I)=TESTONE(I)+TESTONEB(H)
37  CONTINUE
      DO 50 I=1,NPERMS
        IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
50  CONTINUE
      END IF
200 PVALUE=DBLE(NGE)/NPERMS
      RETURN
      END
*
*****
*
*   SUBROUTINE COMPUTES
*
*   THIS SUBROUTINE COMPUTES THE P-VALUE
*
*****

      SUBROUTINE COMPUTES(NPERMS,OBTONE,OBTONEB,TESTONE,
+ TESTONEB,PVALUE,NPAIRS,S,T)

      INTEGER I,NGE,NPERMS,NPAIRS,S,T
      DOUBLE PRECISION PERMSTAT(100000),OBTONE,OBTONEB,TESTONE(100000),
+ TESTONEB(100000),OBSTAT,PVALUE

      NGE=0

      OBSTAT=OBTONE+OBTONEB
      DO 37 I=1,NPERMS
        PERMSTAT(I)=TESTONE(I)+TESTONEB(I)
        IF (PERMSTAT(I).GE.OBSTAT) NGE=NGE+1
37  CONTINUE

      PVALUE=DBLE(NGE)/NPERMS
      RETURN
      END

```

RANDOM NUMBER FUNCTIONS

```
*****  
*  
* THE IGUIN() AND RANF() FUNCTIONS WERE OBTAINED FROM THE RANLIB  
* RANDOM NUMBER FORTRAN LIBRARY COMPILED AND WRITTEN BY BARRY W.  
* BROWN AND JAMES LOVATO  
*  
*****
```

APPENDIX C

SIMULATION RESULTS FOR THE D -PW AND D -AK TESTS

Following are the empirical powers obtained from the Monte Carlo simulation study described in Section 4.2 under each situation for each sample size considered. The situations studied are listed in the tables by the corresponding distribution abbreviation and percent censoring. For instance, *Exp, 30* signifies the previously described *Exp* case with 30% censoring. Asterisks in the tables indicate the test with greater power, in that the corresponding confidence interval given at (4.1) does not contain 0. The results given in the tables were obtained assuming a null correlation of 0.5.

Table C.1 Empirical Power Estimates for $n = 5$, $s = 0$, $t = 5$.

	Situation	Nominal $\alpha = .031$		Nominal $\alpha = .062$		Nominal $\alpha = .093$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.004	.004	.042	.046	.052	.056
	<i>Exp, 70</i>	.001	.001	.006	.007	.007	.007
	<i>LL, 30</i>	.012	.012	.050	.060	.058	.064
	<i>LL, 70</i>	.001	.001	.012	.013	.013	.013
H_a True	<i>Exp Sc, 30</i>	.074	.074	.303	.309 *	.317	.321
	<i>Exp Sc, 70</i>	.001	.001	.019	.019	.020	.020
	<i>Exp Loc, 30</i>	.097	.097	.328	.338 *	.345	.354 *
	<i>Exp Loc, 70</i>	.000	.000	.022	.022	.022	.022
	<i>Gen Exp, 30</i>	.002	.002	.051	.060 *	.067	.070
	<i>Gen Exp, 70</i>	.000	.000	.001	.001	.001	.001
	<i>LL Sc, 30</i>	.064	.064	.232	.247 *	.256	.262 *
	<i>LL Sc, 70</i>	.002	.002	.023	.023	.023	.023
	<i>LL Loc, 30</i>	.099	.099	.307	.321 *	.328	.339 *
	<i>LL Loc, 70</i>	.002	.002	.026	.026	.026	.026

Table C.2 Empirical Power Estimates for $n = 5$, $s = 5$, $t = 0$.

	Situation	Nominal $\alpha = .031$		Nominal $\alpha = .062$		Nominal $\alpha = .093$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.006	.006	.036	.032	.044	.039
	<i>Exp, 70</i>	.000	.000	.005	.005	.005	.005
	<i>LL, 30</i>	.012	.012	.046	.043	.054	.050
	<i>LL, 70</i>	.000	.000	.002	.002	.002	.002
H_a True	<i>Exp Sc, 30</i>	.072	.072	.292 *	.282	.317 *	.308
	<i>Exp Sc, 70</i>	.002	.002	.019	.019	.019	.019
	<i>Exp Loc, 30</i>	.095	.095	.322	.320	.339	.339
	<i>Exp Loc, 70</i>	.001	.001	.028	.028	.028	.028
	<i>Gen Exp, 30</i>	.010	.010	.075 *	.064	.085 *	.079
	<i>Gen Exp, 70</i>	.000	.000	.003	.003	.003	.003
	<i>LL Sc, 30</i>	.063	.063	.224	.217	.246	.241
	<i>LL Sc, 70</i>	.002	.002	.018	.018	.018	.018
	<i>LL Loc, 30</i>	.088	.088	.305	.304	.320	.321
	<i>LL Loc, 70</i>	.006	.006	.033	.033	.034	.035

Table C.3 Empirical Power Estimates for $n = 5$, $s = 5$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.011	.010	.046	.049	.095	.090
	<i>Exp, 70</i>	.002	.004	.038	.045	.089	.100
	<i>LL, 30</i>	.009	.008	.041	.044	.094	.092
	<i>LL, 70</i>	.007	.008	.045	.045	.094	.100
H_a True	<i>Exp Sc, 30</i>	.362	.365	.650	.664 *	.794	.801
	<i>Exp Sc, 70</i>	.060	.061	.242	.262 *	.418	.430
	<i>Exp Loc, 30</i>	.285	.300 *	.607	.605	.743	.745
	<i>Exp Loc, 70</i>	.087	.103 *	.342	.365 *	.517	.535 *
	<i>Gen Exp, 30</i>	.050	.049	.153	.157	.265	.270
	<i>Gen Exp, 70</i>	.009	.009	.044	.050	.111	.116
	<i>LL Sc, 30</i>	.181	.193 *	.446	.449	.604	.605
	<i>LL Sc, 70</i>	.051	.053	.213	.232 *	.351	.358
	<i>LL Loc, 30</i>	.211	.214	.485	.486	.632	.638
	<i>LL Loc, 70</i>	.085	.102 *	.352	.375 *	.538	.553 *

Table C.4 Empirical Power Estimates for $n = 5$, $s = 5$, $t = 10$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.013	.017	.042	.053	.101	.119
	<i>Exp, 70</i>	.008	.010	.044	.054	.096	.108
	<i>LL, 30</i>	.011	.013	.054	.058	.092	.108
	<i>LL, 70</i>	.005	.010	.052	.057	.100	.115
H_a True	<i>Exp Sc, 30</i>	.463	.498 *	.710	.750 *	.843	.862 *
	<i>Exp Sc, 70</i>	.111	.129 *	.332	.384 *	.495	.533 *
	<i>Exp Loc, 30</i>	.327	.360 *	.657	.675 *	.789	.797 *
	<i>Exp Loc, 70</i>	.125	.149 *	.410	.447 *	.586	.635 *
	<i>Gen Exp, 30</i>	.047	.056 *	.182	.210 *	.295	.314 *
	<i>Gen Exp, 70</i>	.014	.018 *	.065	.083 *	.125	.141 *
	<i>LL Sc, 30</i>	.233	.252 *	.489	.527 *	.631	.668 *
	<i>LL Sc, 70</i>	.072	.085 *	.260	.309 *	.419	.448 *
	<i>LL Loc, 30</i>	.213	.235 *	.541	.559 *	.707	.726 *
	<i>LL Loc, 70</i>	.130	.155 *	.402	.453 *	.616	.648 *

Table C.5 Empirical Power Estimates for $n = 5$, $s = 5$, $t = 20$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.008	.016	.060	.077	.110	.140
	<i>Exp, 70</i>	.011	.020	.057	.081	.107	.146
	<i>LL, 30</i>	.011	.012	.049	.060	.097	.119
	<i>LL, 70</i>	.017	.022	.058	.077	.112	.144
H_a True	<i>Exp Sc, 30</i>	.499	.582 *	.761	.823 *	.864	.908 *
	<i>Exp Sc, 70</i>	.147	.198 *	.390	.459 *	.545	.615 *
	<i>Exp Loc, 30</i>	.382	.425 *	.718	.753 *	.841	.865 *
	<i>Exp Loc, 70</i>	.161	.220 *	.492	.584 *	.675	.736 *
	<i>Gen Exp, 30</i>	.070	.103 *	.195	.243 *	.304	.372 *
	<i>Gen Exp, 70</i>	.019	.031 *	.071	.094 *	.122	.158 *
	<i>LL Sc, 30</i>	.256	.310 *	.538	.592 *	.689	.745 *
	<i>LL Sc, 70</i>	.106	.147 *	.328	.372 *	.471	.545 *
	<i>LL Loc, 30</i>	.263	.285 *	.588	.612 *	.757	.769 *
	<i>LL Loc, 70</i>	.154	.191 *	.479	.529 *	.662	.732 *

Table C.6 Empirical Power Estimates for $n = 5$, $s = 10$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.006	.006	.041	.035	.087	.074
	<i>Exp, 70</i>	.005	.005	.044	.042	.088	.083
	<i>LL, 30</i>	.013	.011	.063	.060	.106	.098
	<i>LL, 70</i>	.004	.009	.052	.053	.105	.097
H_a True	<i>Exp Sc, 30</i>	.419 *	.397	.733 *	.704	.848 *	.828
	<i>Exp Sc, 70</i>	.121	.118	.349	.340	.501	.503
	<i>Exp Loc, 30</i>	.374	.368	.662	.653	.788 *	.765
	<i>Exp Loc, 70</i>	.159	.173 *	.444	.453	.632	.627
	<i>Gen Exp, 30</i>	.048 *	.037	.163 *	.138	.275 *	.245
	<i>Gen Exp, 70</i>	.003	.005	.052	.048	.101	.094
	<i>LL Sc, 30</i>	.234 *	.213	.501 *	.489	.644 *	.628
	<i>LL Sc, 70</i>	.089	.096	.286	.280	.428	.429
	<i>LL Loc, 30</i>	.275	.281	.553	.555	.682 *	.672
	<i>LL Loc, 70</i>	.187	.198	.462	.468	.632	.623

Table C.7 Empirical Power Estimates for $n = 5$, $s = 10$, $t = 10$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.020	.019	.058	.059	.118	.117
	<i>Exp, 70</i>	.015	.019	.068	.075	.122	.128
	<i>LL, 30</i>	.013	.011	.060	.060	.107	.107
	<i>LL, 70</i>	.008	.010	.052	.059	.097	.103
H_a True	<i>Exp Sc, 30</i>	.548	.559	.810	.819 *	.902	.905
	<i>Exp Sc, 70</i>	.201	.234 *	.473	.496 *	.625	.648 *
	<i>Exp Loc, 30</i>	.457	.469 *	.758	.757	.870	.871
	<i>Exp Loc, 70</i>	.272	.314 *	.603	.632 *	.752	.766
	<i>Gen Exp, 30</i>	.057	.057	.210	.214	.328	.335
	<i>Gen Exp, 70</i>	.028	.028	.089	.092	.138	.147 *
	<i>LL Sc, 30</i>	.294	.303	.568	.577 *	.707	.709
	<i>LL Sc, 70</i>	.136	.160 *	.363	.383 *	.508	.514
	<i>LL Loc, 30</i>	.374	.380	.648	.651	.778	.770
	<i>LL Loc, 70</i>	.263	.290 *	.577	.593 *	.723	.732

Table C.8 Empirical Power Estimates for $n = 5$, $s = 10$, $t = 20$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.008	.011	.048	.054	.090	.105
	<i>Exp, 70</i>	.010	.013	.043	.053	.091	.111
	<i>LL, 30</i>	.014	.014	.058	.065	.114	.127
	<i>LL, 70</i>	.006	.009	.046	.056	.105	.119
H_a True	<i>Exp Sc, 30</i>	.656	.708 *	.884	.908 *	.942	.955 *
	<i>Exp Sc, 70</i>	.240	.292 *	.525	.571 *	.666	.711 *
	<i>Exp Loc, 30</i>	.547	.562 *	.811	.827 *	.902	.908
	<i>Exp Loc, 70</i>	.305	.362 *	.644	.687 *	.814	.836 *
	<i>Gen Exp, 30</i>	.074	.102 *	.215	.245 *	.338	.369 *
	<i>Gen Exp, 70</i>	.023	.027	.080	.095 *	.143	.172 *
	<i>LL Sc, 30</i>	.391	.424 *	.680	.703 *	.802	.822 *
	<i>LL Sc, 70</i>	.166	.208 *	.388	.445 *	.561	.602 *
	<i>LL Loc, 30</i>	.403	.402	.707	.710	.823	.834 *
	<i>LL Loc, 70</i>	.298	.354 *	.646	.682 *	.796	.811 *

Table C.9 Empirical Power Estimates for $n = 5$, $s = 20$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.013	.013	.053	.039	.105	.082
	<i>Exp, 70</i>	.008	.006	.050	.041	.093	.077
	<i>LL, 30</i>	.011	.008	.050	.034	.102	.079
	<i>LL, 70</i>	.013	.012	.051	.043	.100	.080
H_a True	<i>Exp Sc, 30</i>	.534 *	.472	.793 *	.756	.890 *	.860
	<i>Exp Sc, 70</i>	.206 *	.184	.457 *	.415	.596 *	.554
	<i>Exp Loc, 30</i>	.430	.427	.668 *	.649	.787 *	.763
	<i>Exp Loc, 70</i>	.300	.301	.578 *	.552	.706 *	.684
	<i>Gen Exp, 30</i>	.037 *	.022	.149 *	.101	.236 *	.188
	<i>Gen Exp, 70</i>	.010	.009	.053 *	.042	.115 *	.085
	<i>LL Sc, 30</i>	.269 *	.228	.554 *	.499	.710 *	.660
	<i>LL Sc, 70</i>	.146 *	.129	.380 *	.336	.533 *	.486
	<i>LL Loc, 30</i>	.320	.327	.572	.572	.700	.693
	<i>LL Loc, 70</i>	.283	.301 *	.570 *	.539	.698 *	.671

Table C.10 Empirical Power Estimates for $n = 5$, $s = 20$, $t = 10$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.010	.008	.056	.048	.101	.090
	<i>Exp, 70</i>	.009	.011	.064	.057	.119	.107
	<i>LL, 30</i>	.012	.009	.058	.056	.102	.095
	<i>LL, 70</i>	.011	.011	.062	.057	.111	.106
H_a True	<i>Exp Sc, 30</i>	.672 *	.645	.889 *	.874	.940 *	.934
	<i>Exp Sc, 70</i>	.279	.273	.547 *	.532	.675 *	.659
	<i>Exp Loc, 30</i>	.558	.551	.804 *	.790	.893 *	.884
	<i>Exp Loc, 70</i>	.401	.398	.683	.672	.787 *	.774
	<i>Gen Exp, 30</i>	.065 *	.050	.210 *	.175	.340 *	.305
	<i>Gen Exp, 70</i>	.013	.014	.059 *	.050	.117 *	.106
	<i>LL Sc, 30</i>	.376 *	.352	.653 *	.633	.793 *	.758
	<i>LL Sc, 70</i>	.181	.178	.460	.450	.607 *	.595
	<i>LL Loc, 30</i>	.440	.447	.696 *	.686	.797 *	.785
	<i>LL Loc, 70</i>	.364	.385 *	.639	.641	.776 *	.761

Table C.11 Empirical Power Estimates for $n = 5$, $s = 20$, $t = 20$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.010	.012	.051	.055	.116	.116
	<i>Exp, 70</i>	.009	.009	.051	.059	.097	.106
	<i>LL, 30</i>	.008	.007	.041	.040	.096	.095
	<i>LL, 70</i>	.009	.010	.049	.052	.092	.091
H_a True	<i>Exp Sc, 30</i>	.843	.847	.963	.964	.983	.984
	<i>Exp Sc, 70</i>	.349	.373 *	.643	.663 *	.769	.781 *
	<i>Exp Loc, 30</i>	.708	.711	.910	.907	.953	.949
	<i>Exp Loc, 70</i>	.517	.548 *	.790	.788	.887	.887
	<i>Gen Exp, 30</i>	.088	.092	.257	.257	.398	.399
	<i>Gen Exp, 70</i>	.017	.019	.078	.085	.139	.146
	<i>LL Sc, 30</i>	.536	.539	.789	.792	.883	.888
	<i>LL Sc, 70</i>	.240	.255 *	.535	.549 *	.692	.690
	<i>LL Loc, 30</i>	.552	.551	.806	.809	.904	.903
	<i>LL Loc, 70</i>	.516	.536 *	.778	.775	.875	.883

Table C.12 Empirical Power Estimates for $n = 10$, $s = 0$, $t = 5$.

	Situation	Nominal $\alpha = .011$		Nominal $\alpha = .051$		Nominal $\alpha = .101$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.008	.010	.049	.058	.094	.111
	<i>Exp, 70</i>	.001	.001	.015	.016	.053	.060
	<i>LL, 30</i>	.005	.008	.037	.046	.090	.110
	<i>LL, 70</i>	.003	.003	.018	.019	.046	.049
H_a True	<i>Exp Sc, 30</i>	.257	.278 *	.643	.670 *	.799	.827 *
	<i>Exp Sc, 70</i>	.006	.006	.086	.087	.220	.224
	<i>Exp Loc, 30</i>	.255	.269 *	.675	.684 *	.798	.817 *
	<i>Exp Loc, 70</i>	.005	.005	.120	.121	.297	.302 *
	<i>Gen Exp, 30</i>	.024	.031 *	.124	.146 *	.250	.265 *
	<i>Gen Exp, 70</i>	.001	.001	.009	.010	.038	.041
	<i>LL Sc, 30</i>	.142	.150 *	.451	.484 *	.609	.630 *
	<i>LL Sc, 70</i>	.005	.005	.078	.079	.214	.218
	<i>LL Loc, 30</i>	.216	.227 *	.530	.544 *	.687	.703 *
	<i>LL Loc, 70</i>	.003	.003	.126	.127	.311	.311

Table C.13 Empirical Power Estimates for $n = 10$, $s = 5$, $t = 0$.

	Situation	Nominal $\alpha = .011$		Nominal $\alpha = .051$		Nominal $\alpha = .101$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.008	.006	.037	.032	.092	.081
	<i>Exp, 70</i>	.000	.000	.009	.008	.036	.036
	<i>LL, 30</i>	.007	.007	.052	.047	.105	.093
	<i>LL, 70</i>	.001	.001	.011	.010	.051	.050
H_a True	<i>Exp Sc, 30</i>	.236 *	.230	.674 *	.653	.841 *	.819
	<i>Exp Sc, 70</i>	.004	.004	.085	.084	.232	.232
	<i>Exp Loc, 30</i>	.294 *	.286	.691 *	.682	.824 *	.814
	<i>Exp Loc, 70</i>	.003	.003	.115	.117	.286	.286
	<i>Gen Exp, 30</i>	.022	.022	.140 *	.128	.244 *	.224
	<i>Gen Exp, 70</i>	.000	.000	.019	.016	.054	.054
	<i>LL Sc, 30</i>	.172	.167	.482 *	.452	.618 *	.600
	<i>LL Sc, 70</i>	.004	.004	.080	.081	.233	.233
	<i>LL Loc, 30</i>	.218	.219	.556	.555	.696	.695
	<i>LL Loc, 70</i>	.014	.014	.144	.145	.338	.340

Table C.14 Empirical Power Estimates for $n = 10$, $s = 5$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.005	.005	.049	.051	.105	.105
	<i>Exp, 70</i>	.020	.019	.057	.063	.100	.105
	<i>LL, 30</i>	.011	.009	.054	.055	.112	.110
	<i>LL, 70</i>	.018	.020	.050	.052	.102	.105
H_a True	<i>Exp Sc, 30</i>	.584	.592	.849	.849	.920	.921
	<i>Exp Sc, 70</i>	.159	.178 *	.424	.438 *	.599	.601
	<i>Exp Loc, 30</i>	.498	.510 *	.807	.807	.898	.896
	<i>Exp Loc, 70</i>	.245	.261 *	.572	.584 *	.730	.739
	<i>Gen Exp, 30</i>	.064	.062	.197	.201	.315	.312
	<i>Gen Exp, 70</i>	.010	.009	.049	.054 *	.104	.104
	<i>LL Sc, 30</i>	.321	.326	.600	.610	.736	.741
	<i>LL Sc, 70</i>	.122	.134 *	.364	.368	.518	.528 *
	<i>LL Loc, 30</i>	.366	.368	.665	.668	.782	.786
	<i>LL Loc, 70</i>	.223	.255 *	.576	.585	.731	.738

Table C.15 Empirical Power Estimates for $n = 10$, $s = 5$, $t = 10$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.016	.019	.057	.063	.112	.124
	<i>Exp, 70</i>	.012	.015	.047	.055	.092	.101
	<i>LL, 30</i>	.011	.014	.048	.052	.106	.114
	<i>LL, 70</i>	.009	.010	.048	.058	.102	.120
H_a True	<i>Exp Sc, 30</i>	.615	.650 *	.864	.879 *	.936	.950 *
	<i>Exp Sc, 70</i>	.193	.218 *	.472	.508 *	.637	.670 *
	<i>Exp Loc, 30</i>	.571	.588 *	.850	.861 *	.931	.941 *
	<i>Exp Loc, 70</i>	.285	.316 *	.617	.645 *	.760	.780 *
	<i>Gen Exp, 30</i>	.074	.086 *	.221	.243 *	.357	.391 *
	<i>Gen Exp, 70</i>	.017	.019	.069	.083 *	.133	.150 *
	<i>LL Sc, 30</i>	.373	.396 *	.670	.690 *	.794	.799
	<i>LL Sc, 70</i>	.132	.157 *	.364	.395 *	.518	.548 *
	<i>LL Loc, 30</i>	.413	.423	.685	.690	.823	.826
	<i>LL Loc, 70</i>	.264	.285 *	.627	.641 *	.791	.800 *

Table C.16 Empirical Power Estimates for $n = 10$, $s = 5$, $t = 20$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.009	.012	.048	.061	.102	.119
	<i>Exp, 70</i>	.006	.008	.042	.052	.083	.106
	<i>LL, 30</i>	.009	.012	.050	.058	.103	.122
	<i>LL, 70</i>	.007	.008	.052	.069	.101	.123
H_a True	<i>Exp Sc, 30</i>	.675	.739 *	.888	.915 *	.943	.955 *
	<i>Exp Sc, 70</i>	.256	.301 *	.523	.570 *	.672	.705 *
	<i>Exp Loc, 30</i>	.606	.631 *	.852	.861 *	.934	.944 *
	<i>Exp Loc, 70</i>	.307	.356 *	.684	.719 *	.824	.844 *
	<i>Gen Exp, 30</i>	.084	.105 *	.235	.280 *	.347	.386 *
	<i>Gen Exp, 70</i>	.032	.040 *	.089	.103 *	.160	.184 *
	<i>LL Sc, 30</i>	.409	.452 *	.686	.714 *	.821	.845 *
	<i>LL Sc, 70</i>	.153	.175 *	.414	.446 *	.563	.602 *
	<i>LL Loc, 30</i>	.426	.425	.746	.750	.854	.865 *
	<i>LL Loc, 70</i>	.291	.324 *	.653	.684 *	.812	.830 *

Table C.17 Empirical Power Estimates for $n = 10$, $s = 10$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.015	.012	.048	.045	.109	.101
	<i>Exp, 70</i>	.009	.011	.041	.040	.090	.085
	<i>LL, 30</i>	.015	.013	.052	.046	.106	.098
	<i>LL, 70</i>	.007	.007	.054	.046	.095	.092
H_a True	<i>Exp Sc, 30</i>	.648	.644	.876	.869	.942	.938
	<i>Exp Sc, 70</i>	.182	.180	.490	.491	.665 *	.647
	<i>Exp Loc, 30</i>	.571	.567	.826 *	.815	.906 *	.899
	<i>Exp Loc, 70</i>	.344	.345	.653	.661	.796 *	.786
	<i>Gen Exp, 30</i>	.054 *	.044	.210 *	.185	.333 *	.312
	<i>Gen Exp, 70</i>	.014	.013	.061	.054	.117	.117
	<i>LL Sc, 30</i>	.363	.354	.651 *	.641	.769	.763
	<i>LL Sc, 70</i>	.147	.149	.367	.364	.523	.519
	<i>LL Loc, 30</i>	.416	.423	.688	.686	.833 *	.819
	<i>LL Loc, 70</i>	.312	.321	.619	.612	.753	.752

Table C.18 Empirical Power Estimates for $n = 10$, $s = 10$, $t = 10$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.006	.007	.037	.036	.088	.089
	<i>Exp, 70</i>	.022	.023	.062	.064	.105	.111
	<i>LL, 30</i>	.005	.007	.038	.041	.097	.096
	<i>LL, 70</i>	.010	.011	.052	.057	.102	.098
H_a True	<i>Exp Sc, 30</i>	.746	.754	.931	.933	.968	.971
	<i>Exp Sc, 70</i>	.293	.315 *	.582	.592	.733	.745 *
	<i>Exp Loc, 30</i>	.654	.658	.876	.872	.934	.932
	<i>Exp Loc, 70</i>	.404	.431 *	.703	.708	.839	.845
	<i>Gen Exp, 30</i>	.067	.068	.221	.221	.345	.348
	<i>Gen Exp, 70</i>	.019	.019	.070	.074	.136	.139
	<i>LL Sc, 30</i>	.432	.452 *	.718	.725	.834	.832
	<i>LL Sc, 70</i>	.192	.200	.441	.457 *	.603	.605
	<i>LL Loc, 30</i>	.498	.500	.752	.757	.849	.850
	<i>LL Loc, 70</i>	.399	.422 *	.718	.726	.846	.856

Table C.19 Empirical Power Estimates for $n = 10$, $s = 10$, $t = 20$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.013	.015	.050	.054	.089	.099
	<i>Exp, 70</i>	.008	.012	.047	.059	.107	.125
	<i>LL, 30</i>	.009	.010	.043	.049	.097	.109
	<i>LL, 70</i>	.009	.011	.048	.059	.096	.111
H_a True	<i>Exp Sc, 30</i>	.828	.854 *	.953	.964 *	.981	.986 *
	<i>Exp Sc, 70</i>	.328	.372 *	.599	.649 *	.733	.762 *
	<i>Exp Loc, 30</i>	.699	.710	.918	.923	.962	.965
	<i>Exp Loc, 70</i>	.448	.505 *	.781	.800 *	.893	.896
	<i>Gen Exp, 30</i>	.096	.115 *	.254	.290 *	.397	.424 *
	<i>Gen Exp, 70</i>	.020	.022	.067	.080 *	.122	.141 *
	<i>LL Sc, 30</i>	.547	.572 *	.780	.794 *	.870	.876
	<i>LL Sc, 70</i>	.237	.270 *	.519	.549 *	.685	.710 *
	<i>LL Loc, 30</i>	.555	.553	.818	.822	.899	.898
	<i>LL Loc, 70</i>	.435	.476 *	.782	.797 *	.883	.888

Table C.20 Empirical Power Estimates for $n = 10$, $s = 20$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.006	.005	.055	.040	.110	.091
	<i>Exp, 70</i>	.013	.009	.053	.044	.106	.087
	<i>LL, 30</i>	.008	.006	.051	.040	.108	.092
	<i>LL, 70</i>	.013	.010	.042	.037	.089	.076
H_a True	<i>Exp Sc, 30</i>	.706 *	.664	.914 *	.892	.961 *	.947
	<i>Exp Sc, 70</i>	.291	.286	.568 *	.534	.717 *	.684
	<i>Exp Loc, 30</i>	.631	.631	.826	.824	.901 *	.894
	<i>Exp Loc, 70</i>	.412	.416	.693	.692	.806	.800
	<i>Gen Exp, 30</i>	.052 *	.034	.182 *	.149	.312 *	.258
	<i>Gen Exp, 70</i>	.010	.010	.056	.054	.114 *	.096
	<i>LL Sc, 30</i>	.404 *	.372	.694 *	.663	.799 *	.781
	<i>LL Sc, 70</i>	.173	.169	.383 *	.372	.536 *	.506
	<i>LL Loc, 30</i>	.461	.466	.728	.725	.814	.811
	<i>LL Loc, 70</i>	.402	.407	.698	.701	.814	.804

Table C.21 Empirical Power Estimates for $n = 10$, $s = 20$, $t = 10$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.017	.015	.052	.046	.103	.093
	<i>Exp, 70</i>	.004	.004	.047	.034	.091	.089
	<i>LL, 30</i>	.011	.012	.059	.053	.092	.090
	<i>LL, 70</i>	.016	.014	.060	.052	.110	.106
H_a True	<i>Exp Sc, 30</i>	.805 *	.781	.956	.949	.986	.983
	<i>Exp Sc, 70</i>	.346	.362 *	.641	.636	.772 *	.758
	<i>Exp Loc, 30</i>	.711	.715	.899	.898	.943	.940
	<i>Exp Loc, 70</i>	.515	.514	.801	.792	.882	.879
	<i>Gen Exp, 30</i>	.090 *	.078	.264 *	.245	.395 *	.370
	<i>Gen Exp, 70</i>	.019 *	.015	.078 *	.071	.143	.138
	<i>LL Sc, 30</i>	.521 *	.499	.785 *	.766	.862	.857
	<i>LL Sc, 70</i>	.257	.248	.503 *	.492	.636	.628
	<i>LL Loc, 30</i>	.547	.551	.783	.793 *	.880	.878
	<i>LL Loc, 70</i>	.488	.514 *	.769	.772	.882 *	.869

Table C.22 Empirical Power Estimates for $n = 10$, $s = 20$, $t = 20$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.011	.011	.049	.052	.092	.097
	<i>Exp, 70</i>	.003	.004	.047	.047	.102	.105
	<i>LL, 30</i>	.014	.015	.054	.053	.100	.100
	<i>LL, 70</i>	.011	.012	.058	.057	.105	.108
H_a True	<i>Exp Sc, 30</i>	.910	.912	.979	.979	.990	.992
	<i>Exp Sc, 70</i>	.460	.487 *	.739	.752 *	.848	.853
	<i>Exp Loc, 30</i>	.787	.785	.934	.936	.981	.980
	<i>Exp Loc, 70</i>	.634	.657 *	.858	.852	.930	.925
	<i>Gen Exp, 30</i>	.112	.121 *	.290	.294	.421	.428
	<i>Gen Exp, 70</i>	.022	.030 *	.093	.103 *	.169	.178 *
	<i>LL Sc, 30</i>	.633	.637	.852	.854	.923	.931 *
	<i>LL Sc, 70</i>	.344	.366 *	.618	.625	.740	.751 *
	<i>LL Loc, 30</i>	.677	.677	.864	.870	.933	.932
	<i>LL Loc, 70</i>	.600	.615 *	.838 *	.830	.911	.910

Table C.23 Empirical Power Estimates for $n = 25$, $s = 0$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.010	.009	.048	.054	.100	.105
	<i>Exp, 70</i>	.008	.009	.050	.058	.108	.115
	<i>LL, 30</i>	.006	.007	.035	.035	.086	.092
	<i>LL, 70</i>	.013	.013	.048	.054	.096	.102
H_a True	<i>Exp Sc, 30</i>	.885	.893 *	.974	.976	.990	.991
	<i>Exp Sc, 70</i>	.328	.334	.637	.640	.776	.788 *
	<i>Exp Loc, 30</i>	.896	.897	.979	.978	.996	.996
	<i>Exp Loc, 70</i>	.470	.481 *	.798	.801	.890	.890
	<i>Gen Exp, 30</i>	.085	.095 *	.246	.254	.391	.401 *
	<i>Gen Exp, 70</i>	.013	.014	.063	.062	.119	.126 *
	<i>LL Sc, 30</i>	.615	.627 *	.838	.840	.910	.912
	<i>LL Sc, 70</i>	.253	.259	.529	.544 *	.688	.700 *
	<i>LL Loc, 30</i>	.693	.692	.891	.889	.942	.940
	<i>LL Loc, 70</i>	.479	.491 *	.789	.796	.888	.890

Table C.24 Empirical Power Estimates for $n = 25$, $s = 5$, $t = 0$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.007	.006	.051	.048	.100	.098
	<i>Exp, 70</i>	.011	.012	.053	.054	.101	.100
	<i>LL, 30</i>	.006	.005	.031	.030	.104	.099
	<i>LL, 70</i>	.007	.008	.050	.050	.097	.093
H_a True	<i>Exp Sc, 30</i>	.901	.897	.980	.980	.991	.991
	<i>Exp Sc, 70</i>	.323	.315	.694	.688	.809	.808
	<i>Exp Loc, 30</i>	.859	.858	.970	.969	.986	.986
	<i>Exp Loc, 70</i>	.443	.447	.800	.802	.905 *	.899
	<i>Gen Exp, 30</i>	.087 *	.080	.270 *	.263	.420 *	.399
	<i>Gen Exp, 70</i>	.009	.009	.059	.058	.119	.116
	<i>LL Sc, 30</i>	.605	.603	.847 *	.839	.921	.921
	<i>LL Sc, 70</i>	.228	.231	.561	.558	.725	.723
	<i>LL Loc, 30</i>	.687	.688	.877	.879	.939	.938
	<i>LL Loc, 70</i>	.491	.488	.805	.804	.910	.910

Table C.25 Empirical Power Estimates for $n = 25$, $s = 5$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.012	.013	.065	.067	.108	.107
	<i>Exp, 70</i>	.005	.006	.052	.054	.109	.106
	<i>LL, 30</i>	.010	.012	.047	.047	.084	.084
	<i>LL, 70</i>	.010	.009	.050	.051	.106	.109
H_a True	<i>Exp Sc, 30</i>	.933	.936	.990	.990	.997	.998
	<i>Exp Sc, 70</i>	.474	.483	.759	.762	.859	.862
	<i>Exp Loc, 30</i>	.896	.894	.984	.981	.994	.994
	<i>Exp Loc, 70</i>	.695	.698	.892	.893	.950	.950
	<i>Gen Exp, 30</i>	.118	.115	.323	.320	.462	.459
	<i>Gen Exp, 70</i>	.020	.021	.079	.081	.149	.155
	<i>LL Sc, 30</i>	.689	.690	.887	.888	.934	.935
	<i>LL Sc, 70</i>	.321	.326	.594	.604 *	.736	.737
	<i>LL Loc, 30</i>	.757	.759	.921	.921	.960	.962
	<i>LL Loc, 70</i>	.675	.677	.895	.901	.955	.955

Table C.26 Empirical Power Estimates for $n = 25$, $s = 5$, $t = 10$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.006	.007	.037	.038	.090	.093
	<i>Exp, 70</i>	.011	.010	.044	.048	.088	.094
	<i>LL, 30</i>	.009	.012	.062	.061	.117	.126
	<i>LL, 70</i>	.008	.011	.041	.048	.088	.095
H_a True	<i>Exp Sc, 30</i>	.952	.954	.992	.994	.997	.998
	<i>Exp Sc, 70</i>	.514	.536 *	.795	.809 *	.892	.902 *
	<i>Exp Loc, 30</i>	.931	.931	.989	.990	.997	.997
	<i>Exp Loc, 70</i>	.700	.705	.916	.916	.960	.960
	<i>Gen Exp, 30</i>	.134	.145 *	.343	.352 *	.478	.485
	<i>Gen Exp, 70</i>	.020	.023	.079	.082	.136	.145 *
	<i>LL Sc, 30</i>	.713	.723	.900	.907 *	.950	.955 *
	<i>LL Sc, 70</i>	.343	.362 *	.642	.657 *	.787	.800 *
	<i>LL Loc, 30</i>	.788	.784	.934	.932	.974	.974
	<i>LL Loc, 70</i>	.672	.681	.897	.899	.944	.947

Table C.27 Empirical Power Estimates for $n = 25$, $s = 5$, $t = 20$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.009	.009	.039	.041	.091	.106
	<i>Exp, 70</i>	.010	.012	.057	.061	.094	.106
	<i>LL, 30</i>	.010	.011	.040	.045	.087	.098
	<i>LL, 70</i>	.012	.015	.057	.056	.094	.108
H_a True	<i>Exp Sc, 30</i>	.954	.965 *	.990	.990	.994	.995
	<i>Exp Sc, 70</i>	.536	.561 *	.792	.810 *	.899	.905
	<i>Exp Loc, 30</i>	.929	.927	.989	.990	.998	.999
	<i>Exp Loc, 70</i>	.737	.750 *	.933	.935	.964	.965
	<i>Gen Exp, 30</i>	.166	.185 *	.340	.371 *	.473	.498 *
	<i>Gen Exp, 70</i>	.021	.023	.064	.074 *	.133	.148 *
	<i>LL Sc, 30</i>	.730	.752 *	.909	.915 *	.949	.956 *
	<i>LL Sc, 70</i>	.382	.403 *	.680	.699 *	.798	.812 *
	<i>LL Loc, 30</i>	.809	.800	.939	.937	.977	.977
	<i>LL Loc, 70</i>	.702	.726 *	.920	.925	.969	.969

Table C.28 Empirical Power Estimates for $n = 25$, $s = 10$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.013	.013	.055	.055	.104	.102
	<i>Exp, 70</i>	.010	.009	.056	.058	.102	.101
	<i>LL, 30</i>	.011	.010	.050	.050	.095	.095
	<i>LL, 70</i>	.011	.011	.066	.067	.106	.108
H_a True	<i>Exp Sc, 30</i>	.958	.954	.996	.996	.999	.999
	<i>Exp Sc, 70</i>	.511	.522 *	.779	.777	.879	.882
	<i>Exp Loc, 30</i>	.924	.923	.987 *	.981	.994	.993
	<i>Exp Loc, 70</i>	.715	.722	.914 *	.905	.963 *	.954
	<i>Gen Exp, 30</i>	.112 *	.103	.314 *	.294	.456 *	.447
	<i>Gen Exp, 70</i>	.012	.012	.071	.069	.139	.137
	<i>LL Sc, 30</i>	.714	.709	.908 *	.902	.963	.960
	<i>LL Sc, 70</i>	.352	.352	.628	.627	.758	.755
	<i>LL Loc, 30</i>	.753	.760	.919	.919	.967	.964
	<i>LL Loc, 70</i>	.693	.697	.890	.887	.950	.945

Table C.29 Empirical Power Estimates for $n = 25$, $s = 10$, $t = 10$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.005	.005	.047	.050	.096	.100
	<i>Exp, 70</i>	.017	.017	.064	.066	.110	.109
	<i>LL, 30</i>	.013	.012	.053	.051	.103	.107
	<i>LL, 70</i>	.011	.009	.060	.061	.113	.115
H_a True	<i>Exp Sc, 30</i>	.960	.960	.997	.997	.999	.999
	<i>Exp Sc, 70</i>	.568	.576	.815	.817	.898	.898
	<i>Exp Loc, 30</i>	.942	.941	.992	.991	.998	.998
	<i>Exp Loc, 70</i>	.760	.771 *	.933	.931	.970	.970
	<i>Gen Exp, 30</i>	.162	.158	.389	.391	.527	.530
	<i>Gen Exp, 70</i>	.022	.024	.067	.071	.136	.143 *
	<i>LL Sc, 30</i>	.757	.760	.915	.917	.961	.960
	<i>LL Sc, 70</i>	.383	.394 *	.675	.679	.801	.797
	<i>LL Loc, 30</i>	.804	.805	.942	.939	.968	.970
	<i>LL Loc, 70</i>	.734	.733	.936	.936	.978	.975

Table C.30 Empirical Power Estimates for $n = 25$, $s = 10$, $t = 20$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.009	.012	.061	.062	.098	.101
	<i>Exp, 70</i>	.007	.009	.041	.050	.095	.105
	<i>LL, 30</i>	.007	.010	.044	.046	.091	.094
	<i>LL, 70</i>	.010	.010	.042	.046	.100	.109
H_a True	<i>Exp Sc, 30</i>	.981	.987 *	.999	.999	1.00	1.00
	<i>Exp Sc, 70</i>	.615	.634 *	.841	.855 *	.915	.925 *
	<i>Exp Loc, 30</i>	.954	.949	.993	.992	.996	.996
	<i>Exp Loc, 70</i>	.789	.804 *	.955	.958	.981	.982
	<i>Gen Exp, 30</i>	.164	.181 *	.403	.419 *	.536	.554 *
	<i>Gen Exp, 70</i>	.017	.022 *	.081	.091 *	.158	.175 *
	<i>LL Sc, 30</i>	.817	.825 *	.941	.946 *	.970	.973
	<i>LL Sc, 70</i>	.447	.458 *	.701	.712 *	.814	.824 *
	<i>LL Loc, 30</i>	.842	.840	.958	.957	.984	.982
	<i>LL Loc, 70</i>	.769	.774	.929	.928	.965	.966

Table C.31 Empirical Power Estimates for $n = 25$, $s = 20$, $t = 5$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.015	.012	.053	.051	.097	.085
	<i>Exp, 70</i>	.009	.008	.042	.039	.074	.070
	<i>LL, 30</i>	.012	.012	.049	.046	.106	.099
	<i>LL, 70</i>	.010	.010	.052	.051	.099	.094
H_a True	<i>Exp Sc, 30</i>	.955	.959	.992	.992	.998	.997
	<i>Exp Sc, 70</i>	.541	.539	.809	.805	.898	.892
	<i>Exp Loc, 30</i>	.902	.907	.977	.976	.986	.986
	<i>Exp Loc, 70</i>	.739	.734	.907 *	.899	.952	.946
	<i>Gen Exp, 30</i>	.116 *	.100	.314 *	.297	.469 *	.438
	<i>Gen Exp, 70</i>	.014	.011	.067 *	.061	.133 *	.123
	<i>LL Sc, 30</i>	.719 *	.708	.900 *	.892	.951	.946
	<i>LL Sc, 70</i>	.383	.375	.666 *	.655	.790	.782
	<i>LL Loc, 30</i>	.773	.780	.930	.931	.962	.963
	<i>LL Loc, 70</i>	.722	.722	.887	.885	.937 *	.930

Table C.32 Empirical Power Estimates for $n = 25$, $s = 20$, $t = 10$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.008	.009	.056	.052	.113	.110
	<i>Exp, 70</i>	.015	.015	.052	.050	.104	.102
	<i>LL, 30</i>	.009	.008	.047	.048	.095	.087
	<i>LL, 70</i>	.008	.008	.055	.051	.103	.101
H_a True	<i>Exp Sc, 30</i>	.981	.979	.996	.996	.999	.999
	<i>Exp Sc, 70</i>	.608	.607	.842	.847	.928	.924
	<i>Exp Loc, 30</i>	.955	.956	.992	.993	.996	.996
	<i>Exp Loc, 70</i>	.805	.809	.954 *	.947	.977	.974
	<i>Gen Exp, 30</i>	.131 *	.119	.373 *	.350	.519 *	.500
	<i>Gen Exp, 70</i>	.023 *	.019	.088	.089	.163	.158
	<i>LL Sc, 30</i>	.790	.794	.931	.928	.969	.968
	<i>LL Sc, 70</i>	.411	.417	.708	.704	.838	.834
	<i>LL Loc, 30</i>	.825	.833	.956	.957	.981	.981
	<i>LL Loc, 70</i>	.784	.782	.930 *	.924	.965	.961

Table C.33 Empirical Power Estimates for $n = 25$, $s = 20$, $t = 20$.

	Situation	Nominal $\alpha = .01$		Nominal $\alpha = .05$		Nominal $\alpha = .10$	
		<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>	<i>D-PW</i>	<i>D-AK</i>
H_0 True	<i>Exp, 30</i>	.008	.008	.053	.054	.116	.118
	<i>Exp, 70</i>	.006	.007	.051	.055	.119	.117
	<i>LL, 30</i>	.006	.005	.049	.047	.089	.087
	<i>LL, 70</i>	.010	.011	.054	.057	.106	.109
H_a True	<i>Exp Sc, 30</i>	.992	.991	1.00	1.00	1.00	1.00
	<i>Exp Sc, 70</i>	.697	.712 *	.890	.893	.939	.940
	<i>Exp Loc, 30</i>	.975	.976	.996	.996	.999	.999
	<i>Exp Loc, 70</i>	.866	.868	.968	.969	.989	.987
	<i>Gen Exp, 30</i>	.178	.182	.415	.421	.552	.562 *
	<i>Gen Exp, 70</i>	.015	.017	.076	.083 *	.158	.163
	<i>LL Sc, 30</i>	.848	.855	.963	.964	.984	.985
	<i>LL Sc, 70</i>	.507	.516	.783	.787	.886	.887
	<i>LL Loc, 30</i>	.867	.869	.963	.966	.987	.986
	<i>LL Loc, 70</i>	.820	.825	.952	.947	.978	.978

APPENDIX D
SIMULATION RESULTS FOR THE PAIRED *PW* AND *AK* TESTS

Following are the empirical powers obtained from the Monte Carlo simulation study described in Section 5.3 under each situation for each sample size considered. Again, the situations studied are listed in the tables by the corresponding distribution abbreviation and percent censoring. The results given in the tables were obtained assuming a null correlation of 0.5.

Table D.1 Empirical Power Estimates at $\alpha = .10$.

	Situation	$n = 5$		$n = 10$		$n = 25$	
		Paired <i>PW</i>	Paired <i>AK</i>	Paired <i>PW</i>	Paired <i>AK</i>	Paired <i>PW</i>	Paired <i>AK</i>
H_0 True	<i>Exp, 30</i>	.143	.130	.108	.106	.117	.114
	<i>Exp, 70</i>	.110	.093	.114	.108	.099	.097
	<i>LL, 30</i>	.127	.109	.120	.113	.105	.102
	<i>LL, 70</i>	.112	.093	.115	.111	.093	.094
H_a True	<i>Exp Sc, 30</i>	.611	.559	.836	.826	.997	.997
	<i>Exp Sc, 70</i>	.333	.264	.585	.555	.858	.855
	<i>Exp Loc, 30</i>	.657	.616	.830	.823	.987	.986
	<i>Exp Loc, 70</i>	.385	.331	.698	.682	.945	.945
	<i>Gen Exp, 30</i>	.217	.189	.254	.238	.451	.443
	<i>Gen Exp, 70</i>	.090	.065	.124	.113	.132	.127
	<i>LL Sc, 30</i>	.454	.415	.634	.618	.916	.911
	<i>LL Sc, 70</i>	.309	.253	.517	.504	.739	.735
	<i>LL Loc, 30</i>	.576	.546	.726	.708	.934	.933
	<i>LL Loc, 70</i>	.384	.316	.672	.664	.913	.907

REFERENCES

- Akritis, M.G. (1992), "Rank Transform Statistics with Censored Data," *Statistics and Probability Letters*, 13, 209-221.
- Albers, W. (1988), "Combined Rank Tests for Randomly Censored Paired Data," *Journal of the American Statistical Association*, 83, 1159-1162.
- Andersen, P.K., and Borgan, O. (1985), "Counting Process Models for Life History Data: A Review (with Discussion)," *Scandinavian Journal of Statistics*, 12, 97-158.
- Batchelor, J.R., and Hackett, M. (1970), "HL-A Matching in Treatment of Burned Patients with Skin Allografts," *The Lancet*, 2, 581-583.
- Bradley, J.V. (1968), *Distribution-free Statistical Tests*, Englewood Cliffs, NJ: Prentice-Hall.
- Cheng, K.F. (1984), "Asymptotically Nonparametric Tests with Censored Paired Data," *Communications in Statistics*, 13(12), 1453-1470.
- Conover, W.J., and Iman, R.L. (1981), "Rank Transformations as a Bridge Between Parametric and Nonparametric Statistics," *The American Statistician*, 35, 124-129.
- Cuzick, J. (1985), "Asymptotic Properties of Censored Linear Rank Tests," *Annals of Statistics*, 13, 133-141.
- Dabrowska, D.M. (1989), "Rank Tests for Matched Pair Experiments with Censored Data," *Journal of Multivariate Analysis*, 28, 88-114.
- Dabrowska, D.M. (1990), "Signed-Rank Tests for Censored Matched Pairs," *Journal of the American Statistical Association*, 85, 478-485.
- Dwass, M. (1957), "Modified Randomization Tests for Nonparametric Hypotheses," *Annals of Mathematical Statistics*, 28, 181-187.
- Edgington, E.S. (1987), *Randomization Tests*, New York: Dekker, Inc.
- Edgington, E.S., and Gore, A.P. (1986), "Randomization Tests for Censored Survival Distributions," *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 28, 673-681.
- Efron, B. (1967), "The Two-Sample Problem with Censored Data," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 4, 831-853.

- Feller, W. (1966), *An Introduction to Probability Theory and Its Applications*, Vol 2, New York: Wiley and Sons.
- Fisher, R.A. (1935), *Design of Experiments*, Edinburgh: Oliver and Boyd.
- Fleming, T.R., and Harrington, D.P. (1991), *Counting Processes and Survival Analysis*, New York: Wiley and Sons.
- Gehan, E.A. (1965), "A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples," *Biometrika*, 52, 203-223.
- Gill, R.D. (1980), *Censoring and Stochastic Integrals*, Amsterdam: Mathematisch Centrum, Mathematical Centre Tracts 124.
- Good, P. (1989), "Almost Most Powerful Tests Against Composite Alternatives," *Communications in Statistics*, 18(5), 1913-1925.
- Green, B.F. (1977), "A Practical Interactive Program for Randomization Tests of Location," *The American Statistician*, 31, 37-39.
- Hájek, J., and Šidák, Z. (1967), *Theory of Rank Tests*, New York: Academic Press.
- Hoeffding, W. (1951), "A Combinatorial Central Limit Theorem," *Annals of Mathematical Statistics*, 22, 169-192.
- Hoeffding, W. (1952), "The Large-Sample Power of Tests Based on Permutations of Observations," *Annals of Mathematical Statistics*, 23, 169-192.
- Hollander, M., Pledger, G., and Lin, P. (1974), "Robustness of the Wilcoxon Test to a Certain Dependency Between Samples," *Annals of Statistics*, 2, 177-181.
- Janssen, A., and Brenner, S. (1991), "Monte Carlo Results for Conditional Survival Tests Under Randomly Censored Data," *Journal of Statistical Computation and Simulation*, 39, 47-62.
- Kalbfleisch, J.D. and Prentice, R.L. (1980), *The Statistical Analysis of Failure Time Data*, New York: Wiley and Sons.
- Kaplan, E.L., and Meier, P. (1958), "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, 53, 457-481.
- Kempthorne, O., and Doerfler, T.E. (1969), "The Behaviour of Some Significance Tests Under Experimental Randomization," *Biometrika*, 56, 231-248.
- Lam, F.C., and Longnecker, M.T. (1983), "A Modified Wilcoxon Rank Sum Test for Paired Data," *Biometrika*, 70, 510-513.
- Lamport, L. (1986), *Latex: A Document Preparation System*, Reading, MA: Addison-Wesley.

- Lee, E.T. (1992), *Statistical Methods for Survival Data Analysis*, Second Edition, New York: Wiley and Sons.
- Lehmann, E.L. (1986), *Testing Statistical Hypotheses*, Second Edition, New York: Wiley and Sons.
- Lehmann, E.L., and Stein C. (1949), "On the Theory of Some Nonparametric Hypotheses," *Annals of Mathematical Statistics*, 20, 28-45.
- Mantel, N. (1966), "Evaluation of Survival Data and Two New Rank Order Statistics Arising in its Consideration," *Cancer Chemotherapy Reports*, 50, 163-170.
- Mehta, C.R., Patel, N.R., and Tsaitis, A.A. (1984), "Exact Significance Testing with Ordered Categorical Data," *Biometrics*, 40, 819-825.
- Mehta, C.R., Patel, N.R., and Wei, L.J. (1988), "Constructing Exact Significance Tests with Restricted Randomization Rules," *Biometrika*, 75, 295-302.
- Noether, G.E. (1949), "On a Theorem by Wald and Wolfowitz," *Annals of Mathematical Statistics*, 20, 455-458.
- Noreen, E.W. (1989), *Computer-Intensive Methods for Testing Hypotheses: An Introduction*, New York: Wiley and Sons.
- O'Brien, P.C., and Fleming, T.R. (1987), "A Paired Prentice-Wilcoxon Test for Censored Paired Data," *Biometrics*, 43, 169-180.
- Oden, N.L. (1991), "Allocation of Effort in Monte Carlo Simulation for Power of Permutation Tests," *Journal of the American Statistical Association*, 86, 1074-1076.
- Padgett, W.J., and Wei, L.J. (1983), "Conditionally Distribution-free Test for Censored Bivariate Observations," *Metron*, 41(3), 23-29.
- Pagano, M., and Tritchler, D. (1983), "On Obtaining Permutation Distributions in Polynomial Time," *Journal of the American Statistical Association*, 78, 435-440.
- Pearson, E.S. (1937), "Some Aspects of the Problem of Randomization," *Biometrika*, 29, 53-64.
- Peto, R., and Peto, J. (1972), "Asymptotically Efficient Rank Invariant Test Procedures," *Journal of the Royal Statistical Society A*, 135, 185-206.
- Pitman, E.J.G. (1937a), "Significance Tests Which May Be Applied to Samples from Any Populations," *Journal of the Royal Statistical Society B*, 4, 119-130.
- Pitman, E.J.G. (1937b), "Significance Tests Which May Be Applied to Samples from Any Populations. II. The Correlation Coefficient," *Journal of the Royal Statistical Society B*, 4, 225-232.

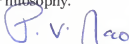
- Pitman, E.J.G. (1938), "Significance Tests Which May Be Applied to Samples from Any Populations. III. The Analysis of Variance Test," *Biometrika*, 29, 322-335.
- Popovich, E.A., and Rao, P.V. (1985), "Conditional Tests for Censored Matched Pairs," *Communications in Statistics*, 14(9), 2041-2056.
- Prentice, R.L. (1978), "Linear Rank Tests with Right Censored Data," *Biometrika*, 65, 167-179.
- Puri, M.L., and Sen, P.K. (1971), *Nonparametric Methods in Multivariate Analysis*, New York: Wiley and Sons.
- Randles, R.H., and Wolfe, D.A. (1991), *Introduction to the Theory of Nonparametric Statistics*, Malabar, FL: Krieger Publishing Company.
- Raychaudhuri, A., and Rao, P.V. (1996), "Efficacies of Some Rank Tests for Censored Bivariate Data," *Nonparametric Statistics*, 6, 1-11.
- Recchia, M., and Rocchetti, M. (1982), "The Simulated Randomization Test," *Computer Programs in Biomedicine*, 15, 111-116.
- Scheffe, H. (1943), "Statistical Inference in the Non-Parametric Case," *Annals of Mathematical Statistics*, 14, 305-332.
- Snijders, T. (1981), "Rank Tests for Bivariate Symmetry," *Annals of Statistics*, 9, 1087-1095.
- Sun, Y., and Sherman, M. (1996), "Some Permutation Tests for Survival Data," *Biometrics*, 52, 87-97.
- Woolson, R.F., and Lachenbruch, P.A. (1980), "Rank Tests for Censored Matched Pairs," *Biometrika*, 67, 597-606.
- Woolson, R.F., and O'Gorman, T.W. (1992), "A Comparison of Several Tests for Censored Paired Data," *Statistics in Medicine*, 11, 193-208.
- Ying, Z., and Wei, L.J. (1994), "The Kaplan-Meier Estimate for Dependent Failure Time Observations," *Journal of Multivariate Analysis*, 50, 17-29.

BIOGRAPHICAL SKETCH

Michael James Dallas was born on October 12, 1970, in Reading, Pennsylvania. After graduating from Wilson High School located in West Lawn, PA, in 1988, he enrolled at Shippensburg University of Pennsylvania. At Shippensburg, Mike majored in mathematics with a computer science concentration, and was also a four year letter winner in baseball, twice earning Academic All-American honors. In May of 1992, he graduated magna cum laude, receiving his Bachelor of Arts degree. He entered graduate school at the University of Florida in August of 1992 and received his Master of Statistics degree in April of 1994. He expects to receive the degree of Doctor of Philosophy in December of 1997. He is a member of Mu Sigma Rho, the American Statistical Association, and an associate member of the Association of General Clinical Research Center Statisticians.

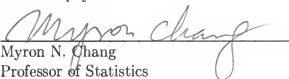
While in graduate school Mike worked as a teaching assistant his first year. For the last four years, he has been working as a consultant for the Division of Biostatistics.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



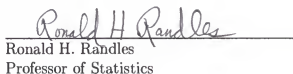
Pejaver V. Rao, Chairman
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Myron N. Chang
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Ronald H. Randles
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Jonathan J. Shuster
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



William M. Mendenhall
Professor of Oncology

This dissertation was submitted to the Graduate Faculty of the Department of Statistics in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 1997

Dean, Graduate School